

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2004, 03:21:30 ; Search time 669 Seconds

(without alignments)  
10172.815 Million cell updates/sec

Title: US-10-031-331B-39

Perfect score: 1602

Sequence: 1 cacaccgtgtattaacat.....aaaaaaaaaaaaaaaaaaaa 1602

Scoring table: IDENTITY NUC

Gapop 10\_0, Gapext 1.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*

1: Geneseqn1980s:\*

2: Geneseqn1990s:\*

3: Geneseqn2000s:\*

4: Geneseqn2001as:\*

5: Geneseqn2001bs:\*

6: Geneseqn2002s:\*

7: Geneseqn2003as:\*

8: Geneseqn2003bs:\*

9: Geneseqn2003cs:\*

10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1602	100.0	1602	4	AAF74206 DNA encod
2	1138.6	71.1	2235	5	AAD16797 Spinach P
3	676.2	42.2	1962	4	AAD44262 Physcomit
4	621.2	38.8	1120	5	AAD16798 Spinach P
5	350.8	21.9	636	8	ACL18332 DNA clone
6	324.6	20.3	605	8	ACL25710 DNA clone
7	304.2	19.0	566	8	ACL19398 DNA clone
8	303.6	19.0	487	9	AD82005 Arabidops
9	302.8	18.9	611	8	ACL25711 DNA clone
10	299.8	18.7	777	8	ACL19405 DNA clone
11	280.6	17.5	575	8	ACL18347 DNA clone
12	276.6	17.3	667	8	ACL18333 DNA clone
13	271.8	17.0	622	8	ACL18325 DNA clone
14	268.6	16.8	589	8	ACL25708 DNA clone
15	255.8	16.0	619	4	AAD44246 Physcomit
16	248.4	15.5	650	8	ACL18342 DNA clone
17	248.2	15.5	736	8	ACL18395 DNA clone
18	240.8	15.0	671	8	ACL19401 DNA clone
19	240.2	15.0	738	8	ACL18354 DNA clone
20	240.0	15.0	517	8	ACL18349 DNA clone
21	239.4	14.9	436	8	ACL18323 DNA clone
22	237.2	14.8	712	8	ACL18373 DNA clone
23	235.2	14.7	663	8	ACL19412 DNA clone

C 24	233.4	14.6	648	8	ACL19403	ACL19403 DNA clone
C 25	233.4	14.6	676	8	ACL18366	ACL18366 DNA clone
C 26	231.2	14.4	727	8	ACL18362	ACL18362 DNA clone
C 27	229.8	14.3	477	8	ACL25703	ACL25703 DNA clone
C 28	227	14.2	621	8	ACL18331	ACL18331 DNA clone
C 29	227	14.2	859	8	ACL18389	ACL18389 DNA clone
C 30	224.6	14.0	696	8	ACL18382	ACL18382 DNA clone
C 31	224.2	14.0	649	8	ACL19408	ACL19408 DNA clone
C 32	224.2	14.0	652	8	ACL19407	ACL19407 DNA clone
C 33	223.6	14.0	624	8	ACL19409	ACL19409 DNA clone
C 34	222.2	13.9	623	8	ACL18326	ACL18326 DNA clone
C 35	222.2	13.9	681	8	ACL18374	ACL18374 DNA clone
C 36	221.6	13.8	697	8	ACL18372	ACL18372 DNA clone
C 37	220.2	13.7	692	8	ACL18391	ACL18391 DNA clone
C 38	213	13.3	587	8	ACL18344	ACL18344 DNA clone
C 39	212.8	13.3	527	8	ACL18334	ACL18334 DNA clone
C 40	211.6	13.2	695	8	ACL18380	ACL18380 DNA clone
C 41	211.2	13.2	684	8	ACL18377	ACL18377 DNA clone
C 42	209.6	13.1	631	8	ACL18356	ACL18356 DNA clone
C 43	209.6	13.1	673	8	ACL18378	ACL18378 DNA clone
C 44	204.8	12.8	617	8	ACL19410	ACL19410 DNA clone
C 45	203.2	12.7	683	8	ACL18383	ACL18383 DNA clone

## ALIGNMENTS

## RESULT 1

AAF74206

ID AAF74206 standard; DNA; 1602 BP.

XX AC AAF74206;

XX DT 06-AUG-2003 (revised)

DT 02-MAY-2001 (first entry)

XX DE DNA encoding environmental stress tolerant protein SEQ ID 39.

XX KM Environmental stress resistance; salt; heat; desert; transgenic plant;

XX KM ds.

XX OS Suaeda japonica.

XX PN WC200106006-A1.

PD 25-JAN-2001.

XX PF 19-JUL-2000; 2000WO-JP004862.

XX PR 19-JUL-1999; 99JP-00235910.

PR 24-MAR-2000; 2000JP-00085377.

XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.

PI Yamada A, Ozeki Y, Saito T;

XX WPI; 2001-147355/15.

XX P-PSDB; AAB80627.

Screening method to obtain DNA encoding environmental stress resistance factor, useful for producing transgenic plants resistant to environmental stress.

Claim 65; Page 123-125; 167pp; Japanese.

Polynucleotide sequences AAF74187 - AAF74218 encode proteins AAB80608 - AAB80639, which impart environmental stress resistance. The invention relates to a method for identifying DNA encoding proteins imparting environmental stress resistance. The method comprises inserting cDNA from a library originating in a salt-resistant organism into a host cell, culturing the transformants under conditions in which the untransformed host does not grow well, and selecting for viable clones. The method is useful for obtaining DNA encoding environmental stress resistance

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2004, 03:21:30 ; Search time 669 Seconds  
(without alignments)  
10172.815 Million cell updates/sec

Title: US-10-031-331B-39

Perfect score: 1602

Sequence: 1 cacaccgttgatttaaccat.....aaaaaaaaaaaaaaaaaaaaa 1602

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 337863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_25Jan04.\*

- 1: Geneseq1980s.\*
- 2: Geneseq1990s.\*
- 3: Geneseq2000s.\*
- 4: Geneseq2001as.\*
- 5: Geneseq2001bs.\*
- 6: Geneseq2002s.\*
- 7: Geneseq2003as.\*
- 8: Geneseq2003bs.\*
- 9: Geneseq2003cs.\*
- 10: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1602	100.0	1602	4	Aaf74206 DNA encod
2	1138.6	71.1	2235	5	Aad16797 Spinach P
3	676.2	42.2	1962	4	Aah44262 Physcomit
4	621.2	38.8	1120	5	Aad16798 Spinach P
5	350.8	21.9	636	8	ACL18332 DNA clone
6	324.6	20.3	605	8	ACL25710 DNA clone
7	304.2	19.0	566	8	ACL19398 DNA clone
8	303.6	19.0	487	9	Ade82005 Arabidops
9	302.8	18.9	611	8	ACL25711 DNA clone
10	299.8	18.7	777	8	ACL19405 DNA clone
11	280.6	17.5	575	8	ACL18347 DNA clone
12	276.6	17.3	667	8	ACL18333 DNA clone
13	271.8	17.0	622	8	ACL18325 DNA clone
14	268.6	16.8	589	8	ACL25708 DNA clone
15	255.8	16.0	619	4	Aah44246 Physcomit
16	248.4	15.5	650	8	ACL18342 DNA clone
17	248.2	15.5	736	8	ACL18395 DNA clone
18	240.8	15.0	671	8	ACL19401 DNA clone
19	240.2	15.0	738	8	ACL18354 DNA clone
20	240.0	15.0	517	8	ACL18349 DNA clone
21	239	14.9	436	8	ACL18323 DNA clone
22	237.2	14.8	712	8	ACL18373 DNA clone
23	235.2	14.7	663	8	ACL19412 DNA clone

C 24	233.4	14.6	648	8	ACL19403	ACL19403 DNA clone
C 25	233.4	14.6	676	8	ACL18366	ACL18366 DNA clone
C 26	231.2	14.4	727	8	ACL18362	ACL18362 DNA clone
C 27	229.8	14.3	477	8	ACL25703	ACL25703 DNA clone
C 28	227	14.2	621	8	ACL18331	ACL18331 DNA clone
C 29	227	14.2	659	8	ACL18389	ACL18389 DNA clone
C 30	224.6	14.0	696	8	ACL18382	ACL18382 DNA clone
C 31	224.2	14.0	649	8	ACL19408	ACL19408 DNA clone
C 32	224.2	14.0	652	8	ACL19407	ACL19407 DNA clone
C 33	223.6	14.0	624	8	ACL19409	ACL19409 DNA clone
C 34	222.2	13.9	623	8	ACL18326	ACL18326 DNA clone
C 35	222.2	13.9	681	8	ACL18374	ACL18374 DNA clone
C 36	221.6	13.8	697	8	ACL18372	ACL18372 DNA clone
C 37	220.2	13.7	692	8	ACL18391	ACL18391 DNA clone
C 38	213	13.3	587	8	ACL18344	ACL18344 DNA clone
C 39	212.8	13.3	527	8	ACL18334	ACL18334 DNA clone
C 40	211.6	13.2	695	8	ACL18380	ACL18380 DNA clone
C 41	211.2	13.2	684	8	ACL18377	ACL18377 DNA clone
C 42	209.6	13.1	631	8	ACL18356	ACL18356 DNA clone
C 43	209.6	13.1	673	8	ACL18378	ACL18378 DNA clone
C 44	204.8	12.8	617	8	ACL19410	ACL19410 DNA clone
C 45	203.2	12.7	683	8	ACL18383	ACL18383 DNA clone

## ALIGNMENTS

## RESULT 1

AAF74206

ID AAF74206 standard; DNA; 1602 BP.

AC AAF74206;

DT 06-AUG-2003 (revised)

DT 02-MAY-2001 (first entry)

XX DNA encoding environmental stress tolerant protein SEQ ID 39.

XX Environmental stress resistance; salt; heat; desert; transgenic plant; ds.

XX Suaeda japonica.

XX WO200106006-A1.

XX 25-JAN-2001.

XX 19-JUL-2000; 2000MO-JP004862.

XX 19-JUL-1999; 99JP-00235910.

XX 24-MAR-2000; 2000JP-00085377.

XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.

XX Yamada A, Ozeki Y, Saito T;

XX WPI; 2001-147355/15.

XX P-PSDB; AAB80627.

XX Screening method to obtain DNA encoding environmental stress resistance factor, useful for producing transgenic plants resistant to environmental stress.

XX Claim 65; Page 123-125; 167pp; Japanese.

XX Polynucleotide sequences AAF74187 - AAF74218 encode proteins AAB80608 - AAB80639, which impart environmental stress resistance. The invention relates to a method for identifying DNA encoding proteins imparting environmental stress resistance. The method comprises inserting cDNA from a library originating in a salt-resistant organism into a host cell, culturing the transformant under conditions in which the untransformed host does not grow well, and selecting for viable clones. The method is useful for obtaining DNA encoding environmental stress resistance



CC factors. The DNA encoding proteins conferring environmental stress  
 CC resistance, can be used in the production of plants resistant to  
 CC environmental stress, which can be cultivated in unfavourable  
 CC environments such as deserts, salt damaged ground, cold regions and the  
 CC oceans. They can be used for increasing the area of land covered by green  
 CC plants, and desert greening and afforestation, in order to counter the  
 CC effects of the increase in atmospheric carbon dioxide concentration. PCR  
 CC primers AAF74219 and AAF74220 are used in an example illustrating the  
 CC method of the invention. (Updated on 06-AUG-2003 to correct OS field.)  
 XX

SQ Sequence 1602 BP; 454 A; 279 C; 383 G; 486 T; 0 U; 0 Other;

Query Match 100.0%; Score 1602; DB 4; Length 1602;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1602; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACACGCTGATTTAAACCATTTGAAGCTATGATGCTCGATTTCTCAAGCTTCTCATCTTGCAC 60  
 DB 1 CACACGCTGATTTAAACCATTTGAAGCTATGATGCTCGATTTCTCAAGCTTCTCATCTTGCAC 60

QY 61 AAGAAGAACGCTGCTGAGATTTCTTCAATGCTTCCGCTCTTGAAGGAAATGCTCTTGG 120  
 DB 61 AAGAAGAACGCTGCTGAGATTTCTTCAATGCTTCCGCTCTTGAAGGAAATGCTCTTGG 120

QY 121 GAACCTGGGCTGCTGATTTGCTGCTTTACTGCTGAATGGCTGAGAAAGCTGCGCAGGTT 180  
 DB 121 GAACCTGGGCTGCTGATTTGCTGCTTTACTGCTGAATGGCTGAGAAAGCTGCGCAGGTT 180

QY 181 ATTGCTCTGGATTTCAATGAGTGTCTATCAAGGAGTCAATGATGCTGAGGCTGCTAC 240  
 DB 181 ATTGCTCTGGATTTCAATGAGTGTCTATCAAGGAGTCAATGATGCTGAGGCTGCTAC 240

QY 241 AAAAAATGTCAGTTTATGCTGCTGATGCTGCTTCTCCACTCTCAGTTTCCACACAT 300  
 DB 241 AAAAAATGTCAGTTTATGCTGCTGATGCTGCTTCTCCACTCTCAGTTTCCACACAT 300

QY 301 TCATTGGATGCTATTTCTCCAAATGGTGTACTCATGTATCTTCTCATGAAGAGGTGAA 360  
 DB 301 TCATTGGATGCTATTTCTCCAAATGGTGTACTCATGTATCTTCTCATGAAGAGGTGAA 360

QY 361 AATTGGGTTGAAGAATGTTGAAATGGTTGAAGCAGGCGGTTACATTTCTTCAGAGAA 420  
 DB 361 AATTGGGTTGAAGAATGTTGAAATGGTTGAAGCAGGCGGTTACATTTCTTCAGAGAA 420

QY 421 TCTTGTCTTCCATCAATCTGGGGATCAAAACGCAAAAGCAATCCACCACTACCGTGAA 480  
 DB 421 TCTTGTCTTCCATCAATCTGGGGATCAAAACGCAAAAGCAATCCACCACTACCGTGAA 480

QY 481 CCTAGGTTCTACACTAAGGCTTCAAGAGTGTCTATTCGAGATGATGCGAATCTCT 540  
 DB 481 CCTAGGTTCTACACTAAGGCTTCAAGAGTGTCTATTCGAGATGATGCGAATCTCT 540

QY 541 TATGAGCTCTCCCTACTAGCTGCAAAATGTTATGGAGCTTATGTCAGAAACAGAAAAAC 600  
 DB 541 TATGAGCTCTCCCTACTAGCTGCAAAATGTTATGGAGCTTATGTCAGAAACAGAAAAAC 600

QY 601 CAGAACCAATAGTTGTTGGTGGCAAAAGTTGATTTCAAGGATGATAGGGGTTCCAG 660  
 DB 601 CAGAACCAATAGTTGTTGGTGGCAAAAGTTGATTTCAAGGATGATAGGGGTTCCAG 660

QY 661 CGATTTCTGGATCTAGCCAGTCAAGTGTATAGCATTTCTGCGATATGAGCGTGTATT 720  
 DB 661 CGATTTCTGGATCTAGCCAGTCAAGTGTATAGCATTTCTGCGATATGAGCGTGTATT 720

QY 721 GGCCTGTTTATGTTAGCACTGGAGGATATGAACCAACAAAGGTTTGTCTCATGCTG 780  
 DB 721 GGCCTGTTTATGTTAGCACTGGAGGATATGAACCAACAAAGGTTTGTCTCATGCTG 780

QY 781 GACTTGAGCTGGCAGAGCTGCTGAGTGTGTTGCTGGAATTTGCTGAGGTTGCTTT 840  
 DB 781 GACTTGAGCTGGCAGAGCTGCTGAGTGTGTTGCTGGAATTTGCTGAGGTTGCTTT 840

QY 841 TACATGCGGAGACCTTTTGAATGTTGAGGTTGTTGATCTCTCGGTTAATATGATT 900

DB 841 TACATGCGGAGACCTTTTGAATGTTGAGGTTGTTGATCTCTCCGTTAATATGATT 900  
 QY 901 TCCCTTGGCCCTTGAGCGTTCTATTTGGGCTTTAAATGTGCTGTTGAGTTTCAGGTAGCAGAT 960  
 DB 901 TCCCTTGGCCCTTGAGCGTTCTATTTGGGCTTTAAATGTGCTGTTGAGTTTCAGGTAGCAGAT 960  
 QY 961 TGCACAAGATTAATTAACCTACCTGATTAACCTTTTGTGATCTATATAGCGCTGACACCAT 1020  
 DB 961 TGCACAAGATTAATTAACCTACCTGATTAACCTTTTGTGATCTATATAGCGCTGACACCAT 1020

QY 1021 CTGCATATTCAGGACAAAGCTGCTGCTGTTTGTAGATCTCTTACAAATGTTGAGCCAGCA 1080  
 DB 1021 CTGCATATTCAGGACAAAGCTGCTGCTGTTTGTAGATCTCTTACAAATGTTGAGCCAGCA 1080

QY 1081 GGTAAAGTTCTAATCAGTGATTACTGCAAGAAAGCTGGTCCACCTCAGCTGAAATTCGCC 1140  
 DB 1081 GGTAAAGTTCTAATCAGTGATTACTGCAAGAAAGCTGGTCCACCTCAGCTGAAATTCGCC 1140

QY 1141 GCTTACATTAAGCAGAGGGGATATGATCTCCATGATGTAAGGATATGCGGAGATGCTT 1200  
 DB 1141 GCTTACATTAAGCAGAGGGGATATGATCTCCATGATGTAAGGATATGCGGAGATGCTT 1200

QY 1201 AAAGATGCTGCAATTTGTTGATGTTCTTCCGAGGATAGAACTCAGCAGTTCTATGCGAGTT 1260  
 DB 1201 AAAGATGCTGCAATTTGTTGATGTTCTTCCGAGGATAGAACTCAGCAGTTCTATGCGAGTT 1260

QY 1261 CTACCGAAGGAACTAGAGACTGTTGAGAGGAAAGGATGTTCTATTAGTGAATTTCTCT 1320  
 DB 1261 CTACCGAAGGAACTAGAGACTGTTGAGAGGAAAGGATGTTCTATTAGTGAATTTCTCT 1320

QY 1321 GAGGAGGATTAACATGACATTTGTTGGAGGTTGGATGATTAAGTTGCGGAGGCTGCCAG 1380  
 DB 1321 GAGGAGGATTAACATGACATTTGTTGGAGGTTGGATGATTAAGTTGCGGAGGCTGCCAG 1380

QY 1381 GGTGAGCAACGATGGGGTCTGTTGTTGCCAAGAAAGTGAAGAATCAGTTGCGGCACT 1440  
 DB 1381 GGTGAGCAACGATGGGGTCTGTTGTTGCCAAGAAAGTGAAGAATCAGTTGCGGCACT 1440

QY 1441 GGCACGTGCTGATTTCTTCTAGTATTAATCTTCAATGTTTTCATGTAATGCTCTACATGT 1500  
 DB 1441 GGCACGTGCTGATTTCTTCTAGTATTAATCTTCAATGTTTTCATGTAATGCTCTACATGT 1500

QY 1501 AAAATTTGCCAATTAAGTTGCTATTTGCGAGACTGTAAGATGATTAATCATATTTATCTTTT 1560  
 DB 1501 AAAATTTGCCAATTAAGTTGCTATTTGCGAGACTGTAAGATGATTAATCATATTTATCTTTT 1560

QY 1561 AATTATCATGATTTATGCAAAAAAATGCAAAAAAATGCAAAAAAATGCAAAAAA 1602  
 DB 1561 AATTATCATGATTTATGCAAAAAAATGCAAAAAAATGCAAAAAAATGCAAAAAA 1602

## RESULT 2

AAD16797  
 ID AAD16797 standard; DNA; 2235 BP.

AC AAD16797;

DT 29-NOV-2001 (first entry)

DE Spinach PEAMT DNA.

XX Spinach; S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase;  
 XX PEAMT; cellular intermediate; phospho-dimethylethanolamine; choline;  
 XX phosphomono-methylethanolamine; phosphocholine; phosphatidylcholine;  
 XX glycine betaine; choline-O-sulphate; lipid content alteration;  
 XX osmotic stress tolerance; nutritional value; transgenic plant;  
 XX cryoprotectant; ds.

OS Spinacia oleracea.

FN Key Location/Qualifiers

FT CDS 259..1738

BEST AVAILABLE COPY





Db 1552 TTGCTGATACATTACGACAGAGGGTTATGATCTCCATAGCTTCAGAGTAGCGGAG 1611  
 Qy 1195 ATGCTTAAGATGCTGATTTGTTGATCTTCCGAGGATAGACATGAGGATTCNTT 1254  
 Db 1612 ATCTGAGATGCTGATTTGTTGATCTTCCGAGGATAGACATGAGGATTCNTT 1671  
 Qy 1255 CGATGCTTACGAGGATGATGATCTTCCGAGGATAGACATGAGGATTCNTT 1314  
 Db 1672 GATGCTTACGAGGATGATGATCTTCCGAGGATAGACATGAGGATTCNTT 1731  
 Qy 1315 TTCTCTGAGGATGATGATCTTCCGAGGATAGACATGAGGATTCNTT 1374  
 Db 1732 TTCTCTGAGGATGATGATCTTCCGAGGATAGACATGAGGATTCNTT 1791  
 Qy 1375 GCCAGGATGATGATCTTCCGAGGATAGACATGAGGATTCNTT 1425  
 Db 1792 TCGATGAGGATGATGATCTTCCGAGGATAGACATGAGGATTCNTT 1842

## RESULT 4

AAD16798  
 ID AAD16798 standard; DNA; 1120 BP.

AC AAD16798;

XX 29-NOV-2001 (first entry)

DT Spinach PEAMT truncated DNA.

XX Spinach; S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase;  
 KW PEAMT; cellular intermediate; phospho-dimethylethanolamine; choline;  
 KW phosphomono-methylethanolamine; phosphocholine; phosphatidylcholine;  
 KW glycine betaine; choline-O-sulphate; lipid content alteration;  
 KW osmotic stress tolerance; nutritional value; transgenic plant;  
 KW cryoprotectant; ds.

XX Spinacia oleracea.

XX Key Location/Qualifiers

PH CDS 254..1120

FT /\*tag= a

FT /product= "Spinach PEAMT truncated protein"

FT /note= "CDS does not include stop codon"

PT /partial

XX WO200168870-A2.

PN 20-SEP-2001.

XX 15-MAR-2001; 2001WO-US008352.

XX 15-MAR-2000; 2000US-00525885.

XX (UYEL) UNIV FLORIDA.

PA (UYCA-) UNIV CARNEGIE MELLON.

XX Hanson AD, Nuccio ML, Henry SA;

PI WPI; 2001-565796/63.

DR P-PSDB; AAE09761.

XX New plant S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase  
 PT polypeptides, useful for modulating the levels of cellular intermediates  
 PT such as phosphodimethylethanolamine and for altering the lipid content in  
 PT plants cells.

XX Claim 18; Page 109; 158pp; English.

XX The present sequence is: spinach S-adenosyl-L-methionine:phospho-  
 CC ethanolamine N-methyltransferase (PEAMT) truncated DNA. The PEAMT  
 CC sequences are useful for modulating the levels of cellular intermediates  
 CC such as phosphodimethylethanolamine, phosphomono-methylethanolamine,

CC choline, phosphocholine, phosphatidylcholine, choline-O-sulphate or  
 CC glycine betaine. They are useful for altering the lipid content in plant  
 CC cells. The polynucleotides are also useful for improving the osmotic  
 CC stress tolerance of a plant and increasing the cryoprotectant properties  
 CC of a plant. The present invention also relates to methods and  
 CC compositions comprising PEAMT used for generating transgenic plants with  
 CC increased nutritional value

XX Sequence 1120 BP; 304 A; 218 C; 256 G; 342 T; 0 U; 0 Other;

Qy Query Match 38.8%; Score 621.2; DB 5; Length 1120;

Best Local Similarity 86.4%; Pred. No. 3.1e-147;

Matches 686; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

Qy 1 CACACCTGCTGATTTAACCATTGAAGCTATGATGCTCGATCTTCAAGCTTCGTGATCTTGAC 60  
 Db 317 CACTCTGTTGATTTGATCTGTTGAGGCTATGATGCTGATTTCAAGCTTCGATCTTGAC 376  
 Qy 61 AAGAAGAACGTCCTGAGATTTCTTCAATGCTTCCGCTCTTTGAAGGAAATGCTCTTG 120  
 Db 377 AAGTGGAGCGACCTGAGGTACTTTCCATGCTTCCACCTTATGAAGGAAATGCTCTTG 436  
 Qy 121 GAACCTGGGCTGATTTGCTGTTTACTGCTGAATGGCTGAGAAAGCTGGCCAGGTT 180  
 Db 437 GAACCTGGGCTGATTTGCTGTTTACTGCTGAATGGCTGAGAAAGCTGGCCAGGTT 496  
 Qy 181 ATTGCTCTGATTTCAATTTGAGAGTCTATCAAGAGAAATGAAGTATCAATGGGCACTAC 240  
 Db 497 ATGCTCTGATTTCAATTTGAGAGTGTATTAAGAGAAATGAAGCATTAATGGGCACTAC 556  
 Qy 241 AAAAAATGCTCAAGTTTATGCTGCTGATGTGACTTCTCCACCTCTCAGTTTCCCAACCAT 300  
 Db 557 AAAAAATGCTCAAGTTTATGCTGCTGATGTGACTTCTCCACCTCTCAGTTTCCCAACCAT 616  
 Qy 301 TCATTGATGATGATTTCTCCATTCGTTTACTGCTGATCTTCTCTGATGAAGAGGTGAA 360  
 Db 617 TCGCTGATGATGATTTCTCCATTCGTTTACTGCTGATCTTCTCTGATGAAGAGGTGAA 676  
 Qy 361 AATTGCTTGAAGAGATTTGAAATGGTTGAAGCCAGGGGGTTACATTTCTCTCAGAGAA 420  
 Db 677 CGTCTGGTTGAAGAGATTTGAAATGGTTGAAGCCAGGGAGGATACATTTCTCTCAGAGAA 736  
 Qy 421 TCTTGTTCATCAATCTGGGATCAAAACGCAAAAGCAATCCCAACCTACCTACCGTCAA 480  
 Db 737 TCTTGTTCATCAATCTGGGATCAAAACGCAAAAGCAATCCCAACCTACCTACCGTCAA 796  
 Qy 481 CTTAGGTTCTACACTAAGGCTTTCBAAGAGTGTCTATTGCAAGATGGATCTGGAACCTCT 540  
 Db 797 CTTAGGTTCTACCAAGATCTTCAAGAGATGGCATATGCAAGATGATTTCTGGGAACCTCC 856  
 Qy 541 TATGAGCTCTCCCTACTTACTGCTCAATGTATTGGAGCTTATGTTCAGAAACAGAAAC 600  
 Db 857 TATGAGCTCTCCCTACTTACTGCTCAATGTATTGGAGCTTATGTTCAGAAACAGAAAC 916  
 Qy 601 CAGAACAGATGATGCTGTTGCTGCAAAAGTTGATTTCTAAGGATGATTAAGGGTTCCAG 660  
 Db 917 CAGAACAGATGATGCTGTTGCTGCAAAAGTTGATTTCTAAGGATGATTAAGGGTTCCAG 976  
 Qy 661 CGATTTCTGGATCTAGCCAGTCAAGTGTATAGCAATCTTCGATATGAGGCTGATTT 720  
 Db 977 CGATTTCTGGATCTAGCTCAATGATTAAGGATTAAGGATTAAGGATTAAGGCTGATTT 1036  
 Qy 721 GGCCTCTGTTATGTTAGCAGTGGAGGATTAAGGATTAAGGATTAAGGATTAAGGCTG 780  
 Db 1037 GGTCTCTGTTATGTTAGTACCGGAGGCTCGAAACCAACCAAGGATTTGTATCAAGCTT 1096  
 Qy 781 GACTTGAAGCCTGG 794  
 Db 1097 GACTTGAAGCCTGG 1110

RESULT 5  
 ACL18332/c



CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley  
 CC varieties, identification of particular varieties and genotype-phenotype  
 CC analysis, isolation of specific genes and creation of new varieties by  
 CC transformation of barley varieties with chem and production of new barley  
 CC varieties with desired properties. The present sequence represents an  
 CC oligonucleotide clone DNA sequence featured in the specification. The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to  
 CC standardise OS field)  
 XX  
 SQ Sequence 605 BP; 186 A; 98 C; 151 G; 170 T; 0 U; 0 Other;

Query Match 20.3%; Score 324.6; DB 8; Length 605;  
 Best Local Similarity 71.1%; Pred. No. 4.4e-72;  
 Matches 429; Conservative 0; Mismatches 174; Indels 0; Gaps 0;

QY 150 TGGTGAATGGCTGAGAAAGCTGGCCAGGTATTTGCTCTGGATTTCATTGAGAGTGTAT 209  
 DB 2 TGGAGAACTGGCTAAGACAGCTGGCATGTTCTTGCAATGGATTTCATTGGAAGTGTAT 61

QY 210 CAAGAAGAAATGAAGTAAATCAATGGGCACTACAAAATGTCAGTTTATGTGTCTGATGT 269  
 DB 62 TAAAAAGATGAAGACATAATGCCATTACGAAATGCATCTTCATGTGTCTGATGT 121

QY 270 GACTTTCGCACTCTGATTTCCACCACTTCATTCATTCATTCATTCATTCATTCATTCAT 329  
 DB 122 TACATCTCCAGACCTGGTGATTCAGGACAACTGGATGATTCATATTTTCAAACTGGCT 181

QY 330 ACTCATGTATCTTCTCATGAAGAGGTGGAAATTTGGTTGAAAGATGTTTGAATGGTT 389  
 DB 182 ACTGATGTATCTTTCAGACGAGGAGGTGCGAAGCTTGTAGAAAGATGTTTAAATGGCT 241

QY 390 GAAGCCAGGGGTTACATTTCTTCAGAGATCTTGTTCATCATCTGCGGATCAAA 449  
 DB 242 CAAGTTGGTGGCATTATCTTTAGAGATCATCTTCCATCATCTGAGACTCAA 301

QY 450 AGCCAAAGCAATCCCAACCACTACCGTGAACCTAGTTCTACATTAAGGCTTCAAAGA 509  
 DB 302 AAGGAAAGTGAATCCGACATATATCGTGAAACCAAGGTTTACACTAAGGTATTTAAGA 361

QY 510 GTGTCATTTGCAAGATGATCTGMAACTCTTATGAGCTCTCCCTACTTACGTCAATG 569  
 DB 362 GGGCCATGCCATGATCAGAGTGGAGCTCTCTGAACTTCTCTGCTTACTTGCAGTG 421

QY 570 TATTGGAGCTTATCTCAGAACAGAACAGACAGACAGATAGTTGGTTGGCAAAA 629  
 DB 422 TGTGGAGCTTATGTGAAGAACAGAAAGATCAAAACCAAGATATGTGGCTATGGCAAAA 481

QY 630 AGTTGATTTCAAGGATGATTAAGGGGTTCCAGCGATTTCTGGATACTAGCCAGTCAAGTG 689  
 DB 482 AGTCAACTCAACAGAGATCGGGGTTTCAAGATTTTGGATATGTGCAAGTCAAAAAC 541

QY 690 TAATAGCACTTCGAGATGAGCGGTATTTGGCCCTGGTTATGTATGTTAGCACTGGAGGATA 749  
 DB 542 CAGTGGATATACGCTATGAGCGGTGTTTGGGCAAGGTTTGGGCAAGCTGGTGGAT 601

QY 750 TGA 752  
 DB 602 TGA 604

## RESULT 7

ACL19398/c  
 ID ACL19398 standard; DNA; 566 BP.  
 XX  
 AC ACL19398;  
 XX  
 XX 17-OCT-2003 (first entry)  
 XX DNA clone originating in barley containing SNP encoding sequence #398.  
 XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;

KW gene; ss.  
 XX  
 OS Hordeum vulgare.  
 XX  
 PN WO2003057877-A1.  
 XX  
 PD 17-JUL-2003.  
 XX  
 PF 16-DEC-2002; 2002WO-1B005403.  
 XX  
 PR 20-DEC-2001; 2001JP-00387059.  
 PR 20-DEC-2001; 2001JP-00387131.  
 PR 20-DEC-2001; 2001JP-00403299.  
 PR 20-DEC-2001; 2001JP-00403300.  
 PR 27-SEP-2002; 2002JP-00327515.  
 XX  
 PA (UYN1-) UNIV JAPAN OKAYAMA.  
 XX  
 PI Sato K, Takeda K, Kohara Y;  
 XX  
 DR WPI; 2003-587127/55.  
 XX  
 PT Single nucleotide polymorphism sites in barley varieties and DNA  
 PT sequences containing them for analysis and identification of barley  
 PT varieties and production of barley transformants with desired  
 PT characteristics.  
 XX  
 PS Disclosure; SEQ ID XX; 284pp; Japanese.  
 XX  
 CC The present invention relates to oligonucleotide clones originating in  
 CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms  
 CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley  
 CC varieties, identification of particular varieties and genotype-phenotype  
 CC analysis, isolation of specific genes and creation of new varieties by  
 CC transformation of barley varieties with chem and production of new barley  
 CC varieties with desired properties. The present sequence represents an  
 CC oligonucleotide clone DNA sequence featured in the specification. The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published-pct-sequences  
 XX  
 SQ Sequence 566 BP; 168 A; 140 C; 97 G; 156 T; 0 U; 5 Other;

Query Match 19.0%; Score 304.2; DB 8; Length 566;  
 Best Local Similarity 71.9%; Pred. No. 6.4e-67;  
 Matches 407; Conservative 0; Mismatches 158; Indels 1; Gaps 1;

QY 485 GGTTCATACCTAAGGCCCTTCAAAGAGTGTCAATTTGCAAGATGATCTCGAAACTCTTATG 544  
 DB 566 GGTTCATACCTANGTATTTAAAGNNGCCATGCCATTGATCAGNAGTGGAGCTCTCTG 507

QY 545 AGCTCTCCCTACTTAGCTGCAAAATGTTTGGAGCTTATGTCAAGAACAGAAACACAGA 604  
 DB 506 AACTTCTCTCTTACTTGTGCAAGTGTGTTGGAGCTTATGTGAAGAACAGAGATCAAA 447

QY 605 ACCAGATTA-GTTGGTTGTGCGCAAAAAGTTGATCTAGGATGATTAAGGGGTTCCAGCGA 663  
 DB 446 ACCAGATATGTTGGCTATGTCGCAAAAAGTCAACTCAACAGAGATCGGGGGTTTCAAAGA 387

QY 664 TTTCTGATACTAGCCAGTACAAAGTGTATATAGCATTTCTGCGATATGAGCGGTATTTGGC 723  
 DB 386 TTTTGGATATGTGCAAGTACAAAACAGTGGAAATATTACGCTATGAGCGGTGTTTTGGG 327

QY 724 CCTGGTTATGTAGCACTGGAGGATATGAACCAACAGAGTTTGTCTCAATGCTGGAC 783  
 DB 326 CAGGTTTTGTGAGCACTGGTGAATTTGACATCAAAAAGAAATTTTGGACTTGTGAT 267

QY 784 TTGAAGCTGCCCAAGAGGCTCTGAAATGTTGGTGTGGAATTTGGAGGTGACATTTTAC 843  
 DB 266 CTTAACCTGGCAGAAAGTGTCTTGTATGTTGGATGGAATCGGGGTGTTGATTTTTC 207

QY 844 ATGGCGAGACCTTTGATGTTGAGTTGTTGATTTGATTTGATTTGATTTGATTTGATTTCC 903



Db 206 ATGGCCGAAACTATGATGTCATGTTGGCAATGATCTTTCATAAACATGTTTCA 147  
 QY 904 TTGGCCCTTGAGCGTTCTTATTTGGGCTTAAATGCTGTGAGTTTCAGGTAGCAGATTGC 963  
 Db 146 TTGGCACTCGAGCATGCTATTGACGCAAGTGGCAGTGGAGTTTGAAGTTGCTGATTGC 87  
 QY 964 ACCAAGATAAACTACCTGTAATCTTTTGTGATCTATATAGCGCTGACACCAATTCG 1023  
 Db 86 ACCACGAGACATACCAGACATATATGATGTTTATCTACAGCGGTGACACCATCTT 27  
 QY 1024 CATATTCAGGACAGCTGCTGGTGTGTT 1049  
 Db 26 CACATACAGATATAACCGCTTTGTT 1

## RESULT 8

ADEB2005

ID ADEB2005 standard; cDNA; 487 BP.

AC ADEB2005;

XX 29-JAN-2004 (first entry)

DE Arabidopsis thaliana expressed polynucleotide seq id 776.

XX genetically modified organism; transgenic organism; plant;  
 KW inhibitor testing; activator testing; modifier testing; fungicide;  
 KW insecticide; genetic function; genetic regulation; cellular metabolism;  
 gene; ss.

XX Arabidopsis thaliana.

OS US2003115639-A1.

XX 19-JUN-2003.

XX 26-JAN-2001; 2001US-00770961.

XX 27-JAN-2000; 2000US-0178466P.

XX (GORL/) GORLACH J.

PA (HAMI/) AN Y.

PA (HAM/) HAMILTON C M.

PA (PRIC/) PRICE J L.

PA (RAIN/) RAINES T M.

PA (YUY/) YU Y.

PA (RAME/) RAMEAKA J G.

PA (PAGE/) PAGE A.

PA (MATH/) MATHW A V.

PA (LEDF/) LEDFORD B L.

PA (WOES/) WOESSNER J P.

PA (HAAS/) HAAS W D.

PA (GARC/) GARCIA C A.

PA (KRIC/) KRICKER M.

PA (SLAT/) SLATER T.

PA (DAVI/) DAVIS K R.

PA (ALLE/) ALLEN K.

PA (HOFF/) HOFFMAN N.

PA (HURB/) HURBAN P.

XX GORLACH J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;

PI Rameaka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;

PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;

CC The invention describes a nucleic acid (I) comprising a sequence capable  
 of hybridising under stringent conditions to any one of 999 fully defined  
 CC Arabidopsis thaliana sequences (1) as given in specification e.g., 360,  
 CC 1137, 455, 219, 472 nucleotides, etc, or its fragment. (I) is useful as a  
 CC hybridisation probe to complementary molecules in a cDNA library. (I) is  
 CC also useful for generating genetically modified and transgenic organisms,  
 CC usually plant cells and plants. A protein encoded by (I) is useful in  
 CC screening assays to determine the effect of candidate inhibitors,  
 CC activators or modifiers of the gene product. The protein is also useful  
 CC for screening biologically active agents e.g., fungicides and  
 CC insecticides. A genetically modified cell, comprising an exogenous  
 CC nucleic acid, where the nucleic acid comprises transcription regulatory  
 CC sequences operably linked to a sequence capable of hybridising under  
 CC stringent conditions to (1) is useful in the study of genetic function  
 CC and regulation, for alteration of the cellular metabolism and for  
 CC screening compounds that may affect the biological function of the gene  
 CC or gene product. This sequence represents an Arabidopsis thaliana  
 CC polynucleotide of the invention.

XX Sequence 487 BP; 129 A; 88 C; 131 G; 139 T; 0 U; 0 Other;

Query Match 19.0%; Score 303.6; DB 9; Length 487;

Best Local Similarity 76.5%; Pred. No. 8.6e-67;

Matches 372; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 543 TGAGCTCTCCCTACTTAGCTGCAAAATGTTATGGAGCTTATGTCAGAAACACAAACCA 602

Db 1 TGAGCTCTCTATGTTGGCTGCAAAATGTTATGGAGCTTATGTCAGAAACACAAACCA 60

QY 603 GAACAGATTAGTTGGTTGGGCAAAAGTTGATTCCTAAGGATGATAGGGGTTCCAGG 662

Db 61 GAATCAGATTGCTGGATATGGCAAAAGTCAAGCTGGAGAAATGCAAGGATTTCCAGG 120

QY 663 ATTTCTGGATAGTACCCAGTACAGTGTATATAGCATTTCTGCGATATGACGGTGTATTTGG 722

Db 121 TTTCTTGGACATGTTCAATACAAAGTCTAGTGGATCTTGGCTATGAGCGTGTCTTTGG 180

QY 723 CCTGGTTATGTTAGCACTGGAGGATATGAAACCAACCAAGATTTGTCTCAATCTCGA 782

Db 181 GGAAGGATATGTGAGCACTGTGTGATTTGAGACACTAAAGAAATTTGTGGCGAAGATGA 240

QY 783 CTTGAAGCTTGGCCAGAAAGTCTCGGATTTGGTGTGGAATTTGGTGGAGGTGACTTTTA 842

Db 241 CCTTAAACCGGACAGAAAGTCTTAGATTTGGTGTGCTATCGGTGGAGGTGACTTTTA 300

QY 843 CATGGCGGAGACCTTTGATTTGAGGTTTGTGATTTGATCTCTCGTTAATATGATTTC 902

Db 301 CATGGCTGAGAAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 360

QY 903 CTTTGGCCCTTGAGCGTTCTATTTGGGCTTAAATGCTGTGTGAGTTTGGAGTGGAGATG 962

Db 361 TTTGCACTGGAGCGGGCCATTGGACTCAATGCTCAGTGGAGTTTGAAGTGGCTGATTG 420

QY 963 CACCAAGATAAACTACCTGATAACTCTTTTGTATGTCATCTATAGCCGTGACCACTTCT 1022

Db 421 CACCAACCAAAACATATCCCGATTAATCTTTTGTATGTCATCTATAGCCGTGACCACTTCT 480

QY 1023 GCATAT 1028

Db 481 GCATAT 486

## RESULT 9

ACL25711

ID ACL25711 standard; DNA; 611 BP.

XX ACL25711;

AC ACL25711;

XX DT 27-OCT-2003 (revised)

XX DT 17-QCT-2003 (first entry)

XX DE DNA clone originating in barley containing SNP encoding sequence #15702.

XX











QY 421 TCTGTTTCCATCATCTGGGATCAAAACGAAAGCAATCCACCACCTACCGTGAA 480  
 |||||  
 Db 509 TCATGCTCCATCAATCTGGAGCTCANAAGGAAAGTGATCCGACACATATATCGTGA 568  
 |||||  
 QY 481 CCTAGGTTCTACACTAGG 499  
 |||||  
 Db 569 CCAAGGTTTACACTAGG 587  
 |||||

## RESULT 15

AAH44246  
 ID AAH44246 standard; DNA; 619 BP.

XX  
 AC AAH44246;

DT 21-SEP-2001 (first entry)

XX Physcomitrella patens 78\_ppprot1\_092\_el2rev gene.

XX Tocopherol and carotenoid metabolism related protein; TCNRP; synthesis;  
 KW Physcomitrella patens; moss; algae; microorganism; fungus; plant;  
 KW identification; genome mapping; modulation; evolutionary study;  
 KW cellular production; fine chemical; ds.

XX Physcomitrella patens.

XX WO200144276-A2.

XX 21-JUN-2001.

PF 14-DEC-2000; 2000WO-EP012698.

PR 16-DEC-1999; 99US-0171121P.

XX (BADI) BASF PLANT SCI GMBH.

XX Lerchl J, Renz A, Ehrhardt T, Reindl A, Cirpus P, Bischoff F;  
 PI Frank M, Freund A, Duwenig E, Schmidt R, Reski R, Badur R;

XX WPI; 2000-398121/42.

DR P-PSDB; AAB99873.

XX Tocopherol and carotenoid metabolism related protein (TCNRP), used to  
 PT produce fine chemicals, is isolated from mosses, algae, microorganisms,  
 PT fungi, plants, or their fragments.

PS Claim 6; Page 107; 123pp; English.

XX The present invention describes isolated tocopherol and carotenoid  
 CC metabolism related proteins (TCNRP) (I) from mosses or algae,  
 CC microorganisms or fungi, plants, or its fragments. (I) can be used as  
 CC enzymes in the production of fine chemicals or in the metabolism of  
 CC tocopherols and carotenoids. (I) also assist in transmembrane transport.  
 CC The fine chemicals that can be produced include lipids, fatty acids,  
 CC vitamins, cofactors, enzymes, amino acids, and nucleotide bases.  
 CC Nucleotide sequences, proteins, vectors and host cells from the present  
 CC invention can be used: (a) to identify mosses related to Physcomitrella  
 CC patens; (b) in mapping genomes of mosses related to Physcomitrella patens  
 CC; (c) in the determination of TCNRP activity; (d) in evolutionary studies;  
 CC (e) in the determination of functional TCNRP regions; (f) and in the  
 CC cellular production of fine chemicals. AAH44222 to AAH44262 encode the  
 CC Physcomitrella patens TCNRP proteins given in AAB99849 to AAB99889.  
 CC AAH44212 to AAH44221 represent nucleotide sequence used in the  
 CC exemplification of the present invention

XX Sequence 619 BP; 176 A; 122 C; 152 G; 169 T; 0 U; 0 Other;

Query Match 16.0%; Score 255.8; DB 4; Length 619;

Best Local Similarity 69.0%; Pred. No. 1.3e-54;

Matches 350; Conservative 0; Mismatches 157; Indels 0; Gaps 0;

QY 919 TCTATTCGGCTTAATGCTGCTTGGTTGAGTTGAGTAGCAGATGCAAGATAACTAC 978  
 |||||

Db 2 TCGATCGCCAGAAAATGTGCACTCGAGTTTGAAGTTGGGATTGCAACCAAGATTAAATTAC 61  
 QY 979 CCTGATAACTCTTTTTCATGTCATCTATAGCGTGACACCATTTCTGCATATTTTCAGACAAG 1038  
 |||||  
 Db 62 CCTCAGCATCTTTTTCATGTCATCTATAGCGTGACACCATTTCTGCATATTTTCAGATAAA 121  
 |||||  
 QY 1039 CTGCGTTGTTTATAGTCTTCTACAAATGGTTGAAGCCAGGAGGTAAAGTTCTTAATCAGT 1098  
 |||||  
 Db 122 CTGCGCTTTTTCACGCTTTTATAAATGGTTGAAGCCTGGAGGTGGGTCGTTATCAGT 181  
 |||||  
 QY 1099 GATTACTGCAAGAAAGCTGCTCCACCCTCACCTGAATTCGCGCTTACATTAAGCAGAGG 1158  
 |||||  
 Db 182 GACTACTGTAGAGCTCCACAACTCGTGGCGGAGTTTCGCTGCATACATTCAGCAGAGG 241  
 |||||  
 QY 1159 GGATATGATCTCCATGATGTAAAGGATATGCGCAGATGCTTAAGATGCTGGATTTGTT 1218  
 |||||  
 Db 242 GGTATGATCTCCATGAGCTTCAGAAATGACGAGAGATGCTGGAAGATGCGCGTTTGTG 301  
 |||||  
 QY 1219 GATGTTCTTCCAGGATAGAACTGAGCAGTTTCATTTCCAGTTCTACGGAAGAACTAGAG 1278  
 |||||  
 Db 302 GAAGTGGTTCGACAGAGACCGCAGATCAGTTTCATTTGAAGTGTACAGAGGAGCTAGCC 361  
 |||||  
 QY 1279 ACTGTTGAGAAAGAAAGGATGTTTCTATTAGTATTTCTCTGAGGAGGATTACAATGAC 1338  
 |||||  
 Db 362 ACCACTGAGCAGCTCGTGACCACTTCATCAGATTTCTCGAGGAGGATTATACTAC 421  
 |||||  
 QY 1339 ATTGTTGAGGTTGGAATGATTAAGTTGCGGAGAGCTGCGCAAGGTTGAGCAAGGATGGGGT 1398  
 |||||  
 Db 422 ATTGTGAGCGGATGGAAGAGTAAAGCTGAAAGCGCTGTTCGAAATGACGAACAGAGTGCGGA 481  
 |||||  
 QY 1399 CTGTTGTTGCCAAGAAAGAGTGAAGA 1425  
 |||||  
 Db 482 CTCTTCATAGCTACAGGCAATTATGA 508  
 |||||

Search completed: August 2, 2004, 05:15:17

Job time : 671 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 2, 2004, 00:19:09 ; Search time 3581 Seconds  
(without alignments)  
3944.376 Million cell updates/sec

Title: US-10-031-331B-40  
Perfect score: 2515  
Sequence: 1 HTVDLTIEMWLDQSASLD.....KLRTAKGEQWGLFVAKKX 473

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 27513289 seqs, 14931090276 residues  
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -MODEL=frame\_p2n\_model -DEV=xlh  
-Q/cg2\_1/USPTO\_spool/US10031331/runat\_26072004.120351.14986/app\_query.fasta\_1.647  
-DB=EST -QPMI=fastap -SUFFIX=first -MINMATCH=0.1 -LOOPCL=0 -LCOREXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=DIOSUM62 -TRANSHUMAN40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10031331@CGN 1.1.2810 @runat\_26072004.120351.14986 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6  
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: EST.\*  
1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estin.\*  
4: em\_estmu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_htc.\*  
9: gb\_est1.\*  
10: gb\_est2.\*  
11: gb\_htc.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: em\_gss\_hum.\*  
18: em\_gss\_inv.\*  
19: em\_gss\_pin.\*  
20: em\_gss\_vrt.\*  
21: em\_gss\_fun.\*  
22: em\_gss\_nam.\*  
23: em\_gss\_mus.\*  
24: em\_gss\_pro.\*  
25: em\_gss\_rtd.\*  
26: em\_gss\_pbg.\*  
27: em\_gss\_vrl.\*  
28: gb\_gssl.\*

29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1937	77.0	2333	11	AY103779	AY103779 Zea mays
2	1420	56.5	1012	11	AY105261	AY105261 Zea mays
3	1413	56.2	1211	10	BH574969	BH574969 E07 Triph
4	1203	47.8	1022	14	CK206655	CK206655 FGAS01825
5	1183	47.0	713	10	BE231445	BE231445 SGS0815 S
6	1162	46.2	767	12	BM408047	BM408047 EST582374
7	1158.5	46.1	1115	14	CK166076	CK166076 FGAS05012
8	1111	44.2	869	14	CF513471	CF513471 CABud0007
9	1106	44.0	781	14	CF208434	CF208434 CAB20003
10	1092	43.4	1127	10	BE420987	BE420987 HM004.F0
11	1066	42.4	822	10	BE040284	BE040284 OE03A05 O
12	1065.5	42.4	733	13	BQ965516	BQ965516 QHB22C02
13	1057	42.0	706	13	BQ120291	BQ120291 EST605867
14	1055	41.9	805	10	BE040460	BE040460 OE06G12 O
15	1038	41.3	721	14	CF482195	CF482195 POL1.5 B0
16	1029	40.9	941	14	CK265554	CK265554 EST711632
17	1027	40.8	770	14	CD320161	CD320161 G608.116A
18	1025	40.8	753	14	CB970167	CB970167 CAB10003
19	1024	40.7	718	14	CA174643	CA174643 SCJFS101
20	1024	40.7	720	14	CF508254	CF508254 USDA-PP-1
21	1018	40.5	731	13	BQ855322	BQ855322 QGB25N08
22	1012.5	40.3	772	9	AJ611652	AJ611652 QGB11652
23	1010	40.2	736	13	BQ852610	BQ852610 QGB18H22
24	1005	40.0	750	12	BM406709	BM406709 EST581036
25	1004	39.9	745	13	BQ762263	BQ762263 EB-co1 SQ
26	1002	39.8	682	13	BQ856500	BQ856500 QGB4H12.Y
27	998	39.7	700	14	CA289355	CA289355 SCAGP1801
28	991	39.4	806	14	CF133358	CF133358 FGAS00364
29	984	39.1	737	14	CF133358	CF133358 WHB4357.A
30	969	38.5	689	13	BQ119788	BQ119788 EST605364
31	961	38.2	672	14	CD232049	CD232049 SSI 31 D1
32	957	38.1	852	12	BG343052	BG343052 HVSMEG000
33	950	37.8	621	13	BU635947	BU635947 O43C05 IN
34	950	37.8	788	14	CA295025	CA295025 SCGLV101
35	948	37.7	848	14	CF513352	CF513352 CABud0007
36	943	37.5	689	12	BG907548	BG907548 Tair1160H
37	942	37.5	664	13	BQ871545	BQ871545 QGB12C12
38	941	37.4	731	13	CA100170	CA100170 SCVPCLE604
39	940	37.4	802	12	BJ573651	BJ573651 BJ573651
40	938	37.3	601	12	BM302134	BM302134 MCA047F06
41	938	37.3	654	12	BJ280399	BJ280399 BJ280399
42	938	37.3	771	14	CA277433	CA277433 SCACSD201
43	938	37.3	853	14	CK204209	CK204209 FGAS01274
44	937	37.3	684	14	CA216364	CA216364 SCRLFL402
45	933	37.1	720	12	BJ573369	BJ573369 BJ573369

ALIGNMENTS

RESULT 1  
AY103779  
LOCUS AY103779 2333 bp mRNA linear HTC 16-OCT-2002  
DEFINITION Zea mays P00126548 mRNA sequence.  
ACCESSION AY103779  
VERSION AY103779.1 GI:21206857  
KEYWORDS HTC.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 2333)

**AUTHORS** Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whittitt, M.S.,  
 Arthur, L.W., Hanafey, M., Morgante, M., and Tingey, S.V.  
**TITLE** Maize Mapping Project/DuPont Consensus Sequences for Design of  
 Overgo Probes  
**JOURNAL** Unpublished (2002)  
**REFERENCE** 2 (bases 1 to 2333)  
**AUTHORS** Coe, E.H.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (25-APR-2002) Maize Mapping Project, University of  
 Missouri, Columbia, MO 65211, USA  
**COMMENT** If you are interested in getting corresponding physical clones,  
 these are publicly available from ZmDB and may be found by BLAST  
 searching at MSL, maizeMap.org; ZmDB, www.zmdb.iastate.edu; TIGR,  
 www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the  
 maize cDNA sequences is either Virginia Walbot, Stanford or Pat  
 Schnable, Iowa State, then clones may be requested from ZmDB:  
 www.zmdb.iastate.edu.

**FEATURES**  
 Location/Qualifiers  
 1..2333  
 /organism="Zea mays"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4577"  
 /db\_xref="taxon:4577"  
 /clone\_lib="Maize Mapping Project/DuPont Consensus  
 Library"  
 /note="this sequence is part of a project of EST  
 assemblies resulting from the application of public  
 contigs to seed DuPont contigs; this resource was  
 assembled by DuPont as part of a collaboration for the  
 overgo addressing of BACs in conjunction with the Maize  
 Mapping Project"

## ORIGIN

Alignment Scores:  
 Pred. No.: 4,066-2331 Length: 2333  
 Score: 1937.00 Matches: 356  
 Percent Similarity: 87.08% Conservative: 55  
 Best Local Similarity: 75.42% Mismatches: 61  
 Query Match: 77.02% Indels: 0  
 DB: 11 Gaps: 0

US-10-031-331B-40 (1-473) x AY103779 (1-2333)

QY 1 HisThrValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp 20  
 DB 488 CACTCCAGTGGCTCACTGTCAGTCCATGATGCTCGACTCCGCGCGCGACCTCGAC 547  
 QY 21 LysGluGluArgProGluLeuLeuSerMetLeuProLeuGluGlyLysCysLeuLeu 40  
 DB 548 AAGGAGAGGAGGACCGAGATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 607  
 QY 41 GluLeuGluAlaGlyIleGlyArgPheThrGlyGluLeuAlaGlyLysAlaGlyGlnVal 60  
 DB 608 GAACCTGGGTGGTAATGGACGCTTACTGGGAGATCTGGCAAAAGAGCTGGGACGTT 667  
 QY 61 IleAlaLeuAspPheIleGluSerAlaIleLysLysAsnGluValIleAsnGlyHisTyr 80  
 DB 668 CTGGCCTAGACTTATTGAAGTGTGATTAGAGAGACCAAGCATTAATGGCATCAC 727  
 QY 81 LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProPhe 100  
 DB 728 AAGACATAACCTTCAGGTGGCGGATGATGATCTTAACGACTTGAAGATTGAAGATAAC 787  
 QY 101 SerLeuAspValIlePheSerAsnTrpLeuLeuMetTyrLeuSerAspGluGluValGlu 120  
 DB 788 TCTGTGTGATCTGATATTTTCAAACTGGCTTAATGATATCTTTCAGATGAGGAGTCCAA 847  
 QY 121 AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyGlyTyrIlePhePheArgGlu 140  
 DB 848 AGCTTGTGGGGAATGTAATGATGATTAAGTGGAGGACATATTCTTCTTCTTCTTCT 907  
 QY 141 SerCysPheHisGlnSerGlyAspHisLysArgLysSerAsnProThrHisTyrArgGlu 160

DB 908 TCATGTTTTTCCACCAATCTGGAGATTCCAAAAGGAGGTGAACCCAAACACACATATCGAGAA 967  
 QY 161 ProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAenSer 180  
 DB 968 CCAAGGTTTTATACCAAGTATTTAAAGAGGGCCATTCATTTGATCAAGATGAGGTTGG 1027  
 QY 181 TyrGluLeuSerLeuLeuSerCysGlyCysGlyIleGlyAlaTyrValArgAsnLysLysAsn 200  
 DB 1028 TTGAACTTTCTCTAGTCACCTGTAAATGCAATTTGGGGCTTATGTCAAAACCAAGAGAA 1087  
 QY 201 GlnAsnGlnIleSerTrpLeuTrpGlnLysValAspSerLysAspLysGlyPheGln 220  
 DB 1088 CAAACCCAGATATGCTGTTATGGAAGAGGTAAATCAACAGACACAGAGATTTTCAA 1147  
 QY 221 ArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArgValPhe 240  
 DB 1148 AGATTCCTGGACACCTGCAATACAAACAGTGGGATATTACGTTATGAGCGTGTCTTT 1207  
 QY 241 GlyProGlyTyrValSerThrGlyGlyTyrGluThrThrLysGluPheValSerMetLeu 260  
 DB 1208 GGTGAAGGTTTGTGAGCAGCTGGTGGATCGAGACTACAAAGGAATTTGTGGGCGATGCTC 1267  
 QY 261 AspLeuLysProGlyGlnLysValLeuAspValGlyCysGlyIleGlyGlyLysPhe 280  
 DB 1268 GATCTTAAACCGGCGCAGAAAGTACTTGTGATGTCGATGTGAATTCGAGCGCGGCGCTT 1327  
 QY 281 TyrMetAlaGluThrPheAspValGluValGlyPheAspLeuSerValAsnMetIle 300  
 DB 1328 TACATGGCTGCAACTATGATGATGCTCATGTTCTTGGTATTGATCTTTGGGTGAACATGTT 1387  
 QY 301 SerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluValAlaAsp 320  
 DB 1388 TCATTTGCAATTGAACGTGCCATTTGACGCGAGTGTCTGTGAATTCGAAGTGTGAT 1447  
 QY 321 CysThrLysIleAsnTyrProAspAsnSerPheAspValIleTyrSerArgSerPhe 340  
 DB 1448 TGCACCAACAAAGGATTACCCAGAAATAGTTTTCAGTCTATCTACAGCGTGCACCATC 1507  
 QY 341 LeuHisIleGlnAspLysProAlaLeuPheArgSerPheTyrLysTrpLeuLysProGly 360  
 DB 1508 CTTTCATACAAAGCAAGCGTCTCTGTTTCAGAGCTTCTTCAATGGGTGAAGCGCGC 1567  
 QY 361 GlyLysValLeuLeuSerAspTyrCysLysLysAlaGlyProSerProGluPheAla 380  
 DB 1568 GGCAAAGTCTTAATCAGCGACTACTGTGAAGATCTTGGAAACCATCAGACAGATTTGCT 1627  
 QY 381 AlaTyrIleLysGlnArgGlyTyrAspLeuHisAspValLysGluTyrGlyGlnMetLeu 400  
 DB 1628 GCGTACATTAAGCAGAGAGGCTATGACCTTCAGCAGCTGAGGCTTATGACAGATGCTG 1687  
 QY 401 LysAspAlaGlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIleArgVal 420  
 DB 1688 AAGGATGCTGGTTTTTCATAATGTCATCGGAAAGATCGCACTGAGCAGTTCTTGAATGT 1747  
 QY 421 LeuArgLysGluLeuGluThrValGluLysGluLysAspValPheIleSerAspPheSer 440  
 DB 1748 CTACAGAGGAGATAGGTGAGTTGAAAGAAACCAAGACGCTTCTCGCAGACTTAC 1807  
 QY 441 GluGluAspTyrAsnAspIleValGlyGlyTrpAsnAspLysLeuArgThrAlaLys 460  
 DB 1808 CAGGAGGATATGACGACATTTGTAATGGTGGACCGGAGCTGAACGGAGCTCTGCC 1867  
 QY 461 GlyGluGlnArgTrpGlyLeuPheValAlaLysLys 472  
 DB 1868 GCGGAGCAGAGGTGGGGTGTGTTTCATTGCCACCAAG 1903

## RESULT 2

AY105261  
 LOCUS Zea.mays.PC012977 mRNA sequence. linear HTC 16-OCT-2002  
 DEFINITION AY105261  
 ACCESSION AY105261.1 GI:21208339  
 VERSION  
 KEYWORDS HTC.





## FEATURES

Location/Qualifiers  
 1. .1211  
 /organism="Triphysaria versicolor"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:64093"  
 /tissue\_type="root-tips -5mm length"  
 /dev\_stage="3-4 weeks growth"  
 /lab\_hosts="E. coli"  
 /clone\_lib="Triphysaria versicolor root-tip, early  
 DMBQ-induced transcript cDNA library"  
 /notes=Vector: pCR2.1 TA Cloning System, Invitrogen,  
 Carlsbad, CA; PCR-based suppression subtractive  
 hybridization cDNA library"

ORIGIN

Alignment Scores:  
 Pred. No.: 9,22e-166 Length: 1211  
 Score: 1413.00 Matches: 260  
 Percent Similarity: 90.25% Conservative: 27  
 Best Local Similarity: 81.76% Mismatches: 31  
 Query Match: 56.18% Indels: 0  
 DB: 1.0 Gaps: 0

US-10-031-331B-40 (1-473) x BE574969 (1-1211)

QY 1 HisThrValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAap 20  
 Db 258 CATTCTGTTGACCTGACCTGGAATCGATCGATGTTGATTCGAAGCATCTGATCTGAT 317  
 QY 21 LysGluGluArgProGluIleLeuSerMetLeuProLeuGluGluLysCysLeuLeu 40  
 Db 318 AAGGAAGAAGGCTGAGGTGTTATCTATCTCTCCACCATATGAAGGCAAGTCGCTTCT 377  
 QY 41 GluLeuGlyAlaGlyIleGlyArgPheThrGlyGluLeuAlaGluLysAlaGlyVal 60  
 Db 378 GAACCTGGAGCTGTTATGTCCTGTTTCACTGCTGAATTAGCTAAGAGGCCAAGAGCTT 437  
 QY 61 IleAlaLeuAspPheIleGluSerAlaIleLysLysAsnGluValIleAsnGlyHisTyr 80  
 Db 438 GTAGCTTTAGACTTTCGGAAGCATATTAAGAAGATGAAGCCCTCATGCGCATCAT 497  
 QY 81 LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProHis 100  
 Db 498 AAAATGTCAAATTCATTTGTGCTGATGTCACATCCCAAGCTTGAACCTTCTCTGAAGGA 557  
 QY 101 SerLeuAspValIlePheSerAsnTrpLeuLeuMetTyrLeuSerAspGluValGlu 120  
 Db 558 TCACCTTGACGTGATTTCTCGACTGCTATGATGATCTTTCAGACAATGAGGTGAG 617  
 QY 121 AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyGlyTyrIlePhePheArgGlu 140  
 Db 618 AATCTTGCAGAGAGATGGTCAATGTTGATGATGTTGGTGTCTATATATTTTCAGAGAA 677  
 QY 141 SerCysPheHisGlnSerGlyAspHisLysArgLysSerAsnProThrHisTyrArgGlu 160  
 Db 678 TCATGTTTCCATCATGTCGCTGACCAAAAGAAAGATATATCCGACTCATTTATCGCAA 737  
 QY 161 ProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSer 180  
 Db 738 CCAAGATTTATACCAAGTTGTTCAAGATGCCATATGATGATGATGATGATGATGATGAT 797  
 QY 181 TyrGluLeuSerLeuLeuSerCysLysCysValIleGlyAlaTyrValArgAsnLysLysAsn 200  
 Db 798 TTGACCTGCTCTTATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 857  
 QY 201 GlnAsnGlnIleSerTrpLeuTrpGlnLysValAspSerLysAspAspLysGlyPheGln 220  
 Db 858 CAGATCAGATCTGCTGATTTGGCAAAAGGTTAGCTCAGATCATGACAGAGGATTCAG 917  
 QY 221 ArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArgValPhe 240  
 Db 918 AAATCTTGATATGTCGAATATTAATGATGATGATGATGATGATGATGATGATGATGAT 977

QY 241 GlyProGlyTyrValSerThrGlyTyrGluThrThrLysGluPheValSerMetLeu 260  
 Db 978 GCACAGGCGATATGTCAGCACCGCTGGACTAGAAACAACTAAGAAATTTGACCAACTG 1037  
 QY 261 AspLeuLysProGlyGlnLysValLeuAspValGlyCysGlyIleGlyGlyAspPhe 280  
 Db 1038 GACCTTAAGTCGGCCCAAGAGTCTAGACGTGGGTGGCATTGGAGAGGTGATTTTC 1097  
 QY 281 TyrMetAlaGluThrPheAspValGluValValGlyPheAspLeuSerValAsnMetIle 300  
 Db 1098 TACATGGCTGACGAGCATGATGTTTCATGTTGTTGGCATTCGCTCGGTCAATGATT 1157  
 QY 301 SerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluVal 318  
 Db 1158 TCTTTGCTCTGAGCGTGCATGTTGCTCAGTGTGCTGTGAGTTGAGGTT 1211

## RESULT 4

CK206655

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

ORIGIN

CK206655 1022 bp mRNA linear EST 08-DEC-2003  
 FGAS018258 Triticum aestivum FGAS: Library 5 GATE 7 Triticum  
 aestivum cDNA, mRNA sequence.

CK206655.1 GI:39569045

EST

Triticum aestivum (bread wheat)

Triticum aestivum

Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Poideae; Triticeae; Triticum.

1 (bases 1 to 1022)

Allard, F., Crosby, W.L., Danyluk, J., Rudes, F., Frick, M., Gaudet, D.,  
 Links, M.G., McCarthy, E.L., Monroy, A., Muzak, I., Nilson, D.,  
 Penniket, C., Roach, J.L. and Sarhan, F.

Functional Genomics of Abiotic Stress in Wheat and Canola Crops

Unpublished (2003)

Contact: Wm L Crosby

Bioinformatics

University of Saskatchewan, Department of Computer Science

1C101 Engineering Building, 57 Campus Drive, Saskatoon,  
 Saskatchewan, S7N 5A9, Canada

Tel: 306 966 1769

Fax: 306 966 2033

Email: fgas.est@cs.usask.ca

This sequence is the direct result of the Base calling software  
 Phred (default parameters). It is the raw base calls. To aid in the  
 identification of the high quality insert the software Lucy  
 (default parameters) has been run on this sequence. Lucy identified  
 the region [9,841].

Plate: L5B008 row: C column: 07.

Location/Qualifiers

1. .1022

/organism="Triticum aestivum"

/mol\_type="mRNA"

/db\_xref="taxon:4565"

/clone\_lib="Triticum aestivum FGAS: Library 5 GATE 7"

/notes=Vector: pCMV.SPC016; Crown and developmental stages  
 of spike formation in wheat cultivar Norstar. 4 mRNA  
 populations were combined before constructing the library.  
 The first mRNA population is from 1cm crown sections after  
 30 days of cold acclimation. The second is from 1cm crown  
 sections after 11 days of deacclimation (before  
 deacclimation plants were fully vernalized for 49 days).  
 The third is from different developmental stages of spike  
 formation (5 to 50mm) that still have not emerged from the  
 leaf (dissection required). The last is from different  
 developmental stages of spike and seed formation after  
 having emerged from the leaf (visible). First strand  
 synthesis in this library was done in the presence of  
 methylated dCTP thereby protecting from internal cleavage  
 with NotI."



## Alignment Scores:

Pred. No.: 1,686-139 Length: 1022  
 Score: 1203.00 Matches: 217  
 Percent Similarity: 85.91% Conservative: 39  
 Best Local Similarity: 72.82% Mismatches: 42  
 Query Match: 47.83% Indels: 0  
 DB: 14 Gaps: 0

US-10-031-331B-40 (1-473) x CK206555 (1-1022)

QY 175 AspGlySerGlyAsnSerTrpGluLeuSerLeuLeuSerCysValleGlyAlaTyr 194  
 DB .....  
 QY 195 ValArganLysLysGlnGlnGlnLysSerTrpLeuTrpGlnLysValAspSerLys 214  
 DB .....  
 QY 87 CTGAAGAGCAAGAGAACACAGACCATGATATGCTATGGGAGAGGTCAGATGCACA 146  
 QY 215 AspAspLysGlyPheGlnArgPheLeuAspThrSerGlnTrpLysCysAsnSerLys 234  
 DB .....  
 QY 147 GAAGCAAAAGGCTTCAGAGATTCCTGGACATGTGCAGTACAAATCCATCGGAATCTG 206  
 QY 235 ArgTrpGluArgValPheGlyProGlyTrpValSerThrGlyGlyTrpLysThrLys 254  
 DB .....  
 QY 207 CATTATGAGCGCGTGTGGAGAGGTTATGTGACGACCGGTGGATTTCGACCAACAAA 266  
 QY 255 GluPheValSerMetLeuAspLeuLysProGlyGlnLysValLeuAspValGlyCysGly 274  
 DB .....  
 QY 275 IleGlyGlyAspPheTrpMetAlaGluThrPheAspValGluValLysValGlyPheAsp 294  
 DB .....  
 QY 327 ATCGAGAGGAGTACTTCTACATGGCTGAACTATGATGTCATGCTCTCGGCATCAT 386  
 QY 295 LeuSerValAnMetLysSerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaVal 314  
 DB .....  
 QY 387 CTTTCATCACAATGGTTCGTCGAGTGCAGCGTGCCTCGGCGCTCGTCTCGGT 446  
 QY 315 GluPheGluValAlaAspCysThrLysLysValLeuSerAspTrpAspAsnSerPheAspValle 334  
 DB .....  
 QY 447 GAGTTTGAGTGTGCTGACGACCAACAGCAATACGACAGACACACAGTTGATGTGATC 506  
 QY 335 TyrSerArgAspThrIleLeuHisIleGlnAspLysProAlaLeuPheArgSerPheTyr 354  
 DB .....  
 QY 507 TACAGCGTGACACCTCTCCACATTCAGACAAACCTGCTCTGTTTCAAGAACTCTTC 566  
 QY 355 LysTrpLeuLysProGlyGlyLysValLeuLysSerAspTrpCysLysLysAlaGlyPro 374  
 DB .....  
 QY 567 AAGTGGCTCAAGCGCTGGGGGCAAGTCTGTATCAGTCACTGACGAGGAGCGCTGGGACA 626  
 QY 375 ProSerProGluPheAlaAlaTyrIleLysGlnArgGlyTyrAspLeuHisAspValLys 394  
 DB .....  
 QY 627 CGGTCAAGAGAAATTGCTGCATACATCAAGCAGAGAGGCTATGACCTCCATCAGCTGAAG 686  
 QY 395 GluTrpGlyGlnMetLeuLysAspAlaGlyPheValAspValLeuAlaGluAspArgThr 414  
 DB .....  
 QY 687 ACCTATGGAAGATGCTTGAGATGCGGTTTCAAGATGTGTCGCCGAAGACCGACCC 746  
 QY 415 GluGlnPheIleArgValLeuArgLysGluLeuGluThrValGluLysGluLysAspVal 434  
 DB .....  
 QY 747 GACCACTCTCAGCGTCTCTGAGAGGAGGCTGGCGCGAGAACGAGAGAACAGAGGCGCC 806  
 QY 435 PheIleSerAspPheSerGluAspTrpAsnAspIleValGlyTrpAsnAspLys 454  
 DB .....  
 QY 807 TTCTCGCGGAGCTTCAACGAGAGGACTACGACGACATCTGCAACGCTGGAGCGGAG 866  
 QY 455 LeuArgThrAlaLysGlyGluGlnArgTrpGlyLeuPheValAlaLysLys 472  
 DB .....  
 QY 867 CTCAAGCGAGCTCGCGCGAGCAGAGTGGCGGCTGTTCTATCGCAACAGG 920

## RESULT 5

BE231445 713 bp mRNA linear EST 16-OCT-2000  
 LOCUS

## DEFINITION

SBS0815 Suaeda salsa ZAP cDNA library Suaeda maritima subsp. salsa cDNA similar to c.elegans cDNA YK92B11.3, mRNA sequence.

## ACCESSION

BE231445

## VERSION

BE231445.1

## KEYWORDS

EST.

## SOURCE

ORGANISM

Suaeda maritima subsp. salsa

Suaeda maritima subsp. salsa

Rukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales; Anaranthaceae; Suaeda.

1 (bases 1 to 713)

Zhang, L., Ma, C.L., Wang, P.P., Sun, Y.F., Zhao, Y.X. and Zhang, H.

Expressed sequence tags from a halophyte Suaeda salsa cDNA library

Unpublished (2000)

CONTACT: Hui Zhang

Key Laboratory of Plant Stress Research

The Biology Department of Shandong Normal University

No. 88, Wenhua East Road, Jinan, Shandong Province, 250014, PRC

Tel: (86)531-2960864

Fax: (86)531-2966954

Email: zhanghui@sdnu.edu.cn.

Location/Qualifiers

1..713

source

/organism="Suaeda maritima subsp. salsa"

/mol\_type="mRNA"

/sub\_species="salsa"

/db\_xref="taxon:126914"

/dev\_stage="seedling"

/clone\_lib="Suaeda salsa ZAP cDNA library"

/notes="Organ: aerial part tissue; Vector: lambda zap; Site 1: EcoRI; Site 2: XhoI; total RNA extraction from NaCl(400mM) treated Suaeda salsa by RNeasy kit (Promega); mRNA isolation by MESSAGEMAKER kit (GIBCO BRL); directional cDNA synthesis (Scorei XhoI) by cDNA synthesis kit (STRATAGENE); the ZAP express library by GigapackIII Gold Cloning kit (STRATAGENE)"

ORIGIN

## Alignment Scores:

Pred. No.: 2,996-137 Length: 713

Score: 1183.00 Matches: 231

Percent Similarity: 97.48% Conservative: 1

Best Local Similarity: 97.06% Mismatches: 3

Query Match: 47.04% Indels: 4

DB: 10 Gaps: 0

US-10-031-331B-40 (1-473) x BE231445 (1-713)

QY 81 LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProHis 100  
 DB 1 AAAAAATGTCAGTATTATGTGTGCTGATGTGACTTCTCCCACTCTCAGTTTCCACACNT 60  
 QY 101 SerLeuAspValIlePheSerAsnTrpLeuLeuMetTyrLeuSerAspGluValGlu 120  
 DB 61 TCATTGGATGTGATATTCTCCAAATGGTTACTCATGTATCTTTCTGATGAAGAGGTGNA 120  
 QY 121 AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyTyrIlePhePheArgGlu 140  
 DB 121 GATTTGGTTGAAGAAATGTTGAATGTTGACCCAGGGGGTTATATTTCTTCAGAGNA 180  
 QY 141 SerCysPheHisGlnSerGlyAspHisLysArgLysSerAsnProThrHisTyrArgGlu 160  
 DB 181 TCTTGTTCATCAATCTGGGATCACAAACGCAAAAGCAATCCACCACTTACCGTGA 240  
 QY 161 ProArgPheThrTrpLysAlaPheLysGluCysHisGluAspGlySerGlyAsnSer 180  
 DB 241 CTTAGGTTCTACCTAGAGCCCTTCAAGAGTGTCAATTCGAGATGGATCTCGAAGACTCT 300  
 QY 181 TyrGluLeuSerLeuLeuSerCysLysCysIleGlyAlaTyrValArganLysLysAsn 200  
 DB 301 TATGAGCTCTCCCTACTAGCTGCAATGATTATGGAGCTTATGTCAGAAACCAAGAAAC 360  
 QY 201 GlnAsnGlnIleSerTrpLeuTrpGlnLysValAspSerLysAspAspLysGlyPheGln 220

```

Db      361 CAGAACCAAGATTAGTGGTGTGGCAAAAGATTGATCTTAAGGATGATGAAGGGTCCAG 420
Qy      221 ArgPheLeuAapThrSerGlnTyLysCysAenSerIleLeuArgTyrglu-ArgValph 240
Db      421 CGATTTCGGTACTAGCCAGTACAGTGTGAATAGCAATCTCGCATGATGAAGCGTGTATT 480
Qy      240 eGlyProGlyTyrgValserThrGlyGlyTyrgluThrThrLysGluPheValserMetLe 260
Db      481 TGSCCCTGGTATGTTAGTACCTGGAGGATATGAACCAACCAAGAGTTTGTCTCAATGCT 540
Qy      260 uAepLeuLysPheGlnTyLysValLeuAapValGlyCysGlyIleGlyGlyAspPh 280
Db      541 GACATGAGGCTCC-CHGAGGTCCTGGAGTGTGGTGTGGAAATGGTGGAGGTGACTT 599
Qy      280 eTyMetAlaGluThrPheAapValGluValValGly-PheAapLeuSerValAenMetI 300
Db      600 TTACATGGCGGAGACCTTGTGATGTGGGGTGTGGAAATTTGATCTCTCCGTTAATATGA 659
Qy      300 leSerPheAlaLeu-GluArgSerIleGlyLeuLysCysAlaValGlu 315
Db      660 TTTCCTTTGGCCTTGAGCGGTCTATTGGGCTTAATGTGCTGTGG 707

```

## RESULT 6

```

BM408047
LOCUS      EST582374 potato roots Solanum tuberosum cDNA clone cPRO33G1 5',
end, mRNA sequence.
ACCESSION  BM408047
VERSION     BM408047.1 GI:18259677
KEYWORDS    EST.
SOURCE      Solanum tuberosum (potato)

```

## ORGANISM

```

Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum.
1 (bases 1 to 767)
van der Hoeven,R., Sun,H., Karamycheva,S.A.; Tsai,J., Van Aken,S.,
Utterback,T., Chiemingo,A., Bougri,O., Buell,C.R., Ronning,C.,
Tanksley,S. and Baker,B.
Generation of ESTs from potato roots
Unpublished (2001)
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: T3.

```

## FEATURES

```

source
Location/Qualifiers
1..767
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cPRO33G1"
/tissue_type="roots"
/dev_stage="in vitro grown stem cuttings"
/lab_host="SOLR"
/clone_lib="potato roots"
/notes="vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Cornell University, Tanksley lab;
sequencing: The Institute for Genomic Research. Roots were
isolated from in vitro grown stem cuttings on CM medium.
Roots were isolated two weeks after placing the stem
cuttings from in vitro grown plants on medium."

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 1..47e+134 Length: 767
Score: 1162.00 Matches: 211
Percent Similarity: 89.02% Conservative: 16

```

```

Best Local Similarity: 82.75% Mismatches: 28
Query Match: 46.20% Indels: 0
DB: 12 Gaps: 0
US-10-031-331b-40 (1-473) x BM408047 (1-767)

```

```

Qy      147 GlyAapHisLeuAapGlySerAapProThrHisTyrgluProArgPheTyrgluys 166
Db      1 GAGGACCAAGCAGGAGAGACACCAACCATTTATCGGAGGCTAGATTTTACACAAAG 60
Qy      167 AlaPheLysGluCysHisLeuGlnAapGlySerGlyAenSerTyrgluLeuSerLeuLeu 186
Db      61 GTGTTTAAAGAAATGTCATATAAATGCTGCTGATGCTGATGCTGATGCTGCTGCTCATC 120
Qy      187 SerCysLeuGlyCysIleGlyAlaTyrgluValArgAenLysAenGlnAenGlnIleSerT 206
Db      121 GGTTCGAAGTCATTCGAGCTTAAGTTTAAACCAACCAAGAAATGATGCTGCTGCTGCTGCTG 180
Qy      207 LeuTTCGlnLysValAapSerLysAapSerLysGlyPheGlnArgPheLeuAapThrSer 226
Db      181 CTTATGCAAAAGGTAAATTTCTGAGGATGACAGGGGATTCAGCGTTTCCTGACCAATGTT 240
Qy      227 GlnTyrgLysCysAenSerIleLeuArgTyrgluValPheGlyProGlyTyrgluValSer 246
Db      241 CAATACAAATGTAGTGGCATACTGCCATATGAACGTGCTTTGGAGAAAGTTTATGTGAGC 300
Qy      247 ThrGlyGlyTyrgluThrThrLysGluPheValserMetLeuAapSerProGlyGln 266
Db      301 ACAGAGGACCTTGATACCAACCAAGAAATTCGTTTCTATGTTGATGCTTCACTGGCCAA 360
Qy      267 LysValLeuAapValGlyCysGlyIleGlyGlyAapPheTyrgluMetAlaGluThrPhe 286
Db      361 AAAGTCTTGTGATGTGGTGTGGAAATAGTGGAGGTGACTTTTACATGCTGAGAATAT 420
Qy      287 AspValGluValValGlyPheAapLeuSerValAenMetIleSerPheAlaLeuGluArg 306
Db      421 GATGTTTATGTTGTTGCTATTGACCTTCGATTAACATGATTCGTTGCTGCTGCTGCTG 480
Qy      307 SerIleGlyLeuLysCysAlaValGluPheGluValAlaAapCysThrLysIleAenTy 326
Db      481 TCATTTGCTCAATGTGGTGTGAATTTGAGGTGCTGATTTGTACAAAGAAACATAT 540
Qy      327 ProAapSerPheAapValIleTyrgSerArgAapThrIleLeuHisIleGlnAapLys 346
Db      541 CTTGATGGCAATTTGATGTGATTTACAGTCGAGACACTATCTTCACATCCACACAA 600
Qy      347 ProAlaLeuPheArgSerPheTyrgLysTyrgLysProGlyGlyLysValLeuLys 366
Db      601 CCCAGTTATTACAGATCTTCTACAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Qy      367 AspTyrgLysLysAlaGlyProProSerProGluPheAlaAlaTyrgLysGlnArg 386
Db      661 GATTACTGCAAAATGCCCAATACCAGCATCCGACGAGTTTCTGAGTATATAAAGCAAGG 720
Qy      387 GlyTyrgAapLeuHisAapValLysGluTyrgLysGlnMetLeuLys 401
Db      721 GGTATGATTATCATGATGTTGCAACATATATGCGCAGATGCTCAAA 765

```

## RESULT 7

```

CK166076/c
LOCUS      FGAS050129 Triticum aestivum FGAS: Talt7 Triticum aestivum cDNA,
mRNA sequence.
ACCESSION  CK166076
VERSION     CK166076.1 GI:38998769
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)

```

## ORGANISM

```

Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
1 (bases 1 to 1115)
Allard,F., Crosby,W.L., Danyluk,J., Eudes,F., Frick,M., Gaudet,D.,

```

Genswein, B., Graf, R., Gulick, P., Hrycan, L.D., Laroche, A.,  
 Links, M.G., McCarthy, E.L., Monroy, A., Muzak, I., Nilsson, D.,  
 Penniket, C., Roach, J.L. and Sarhan, F.,  
 Functional Genomics of Abiotic Stress in Wheat and Canola Crops  
 Unpublished (2003)  
 Contact: Wm L Crosby  
 Bioinformatics  
 University of Saskatchewan, Department of Computer Science  
 1C101 Engineering Building, 57 Campus Drive, Saskatoon,  
 Saskatchewan, S7N 5A9, Canada  
 Tel: 306 966 1769  
 Fax: 306 966 2033  
 Email: f9ae.ets@cs.usask.ca

This sequence is the direct result of the Base calling software  
 Phred (default parameters). It is the raw base calls. To aid in the  
 identification of the high quality insert the software Lucy  
 (default parameters) has been run on this sequence. Lucy identified  
 the region [65,863].  
 Plate: TaLc707 row: K column: 02.  
 Location/Qualifiers

## FEATURES

1. .1115  
 /organism="Triticum aestivum"  
 /mol\_type="mRNA"  
 /cultivar="wheat line CI 14106"  
 /db\_xref="taxon:4565"  
 /lab\_host="DMS alpha"  
 /clone\_lib="Triticum aestivum FGAS: TaLc7"  
 /notes="Organ: Crown; Vector: pGEM-T; SSH (suppression  
 subtractive hybridization) cDNA library from genotype  
 CI14106 cold hardened at 2 C for 21 days and 49 days  
 (equal amount of cDNA pooled together before subtraction,  
 tester) and subtracted against genotype CI14106  
 non-hardened (20 C) (driver). Nitro-pyrole anchored  
 oligo-dT priming and non-directional cloning."

## ORIGIN

Alignment Scores:  
 Pred. No.: 7.58e-134 Length: 1115  
 Score: 1158.50 Matches: 225  
 Percent Similarity: 82.66% Conservative: 42  
 Best Local Similarity: 69.66% Mismatches: 56  
 Query Match: 46.06% Indels: 4  
 DB: 14 Gaps: 0

US-10-031-331B-40 (1-473) x CK166076 (1-1115)

QY 150 LysArgLysSerAnProThrHisTyrA:ggluProArgPheTyrThrLysAlaPheLys 169  
 :::  
 DB 1110 CAGAGGGAGAACTACCGCCACTACCGGAGCCGAGGTTTACCACCAAGGTTGTAAG 1051  
 QY 170 GluCyHisLeuGlnAapGlySerGlyAenSerTyrGluLeuSerLeuLeuSerCyLys 189  
 :::  
 DB 1050 GA-TGCCACTCTAGAGCCA-AGAGGGGATTCCTTGAGCTTCCTCGTAACCTCCAG 993  
 QY 190 CysLleGlyAlaTyrValA:gaenLysLysAenGlnAenGlnLleSerTyrLeuTyrGln 209  
 :::  
 DB 992 TGCAAT-GGAGCTTATGTGAAGCAGAGAGAACCCAGACCCAGATATGTGGCTGGGAG 934  
 QY 210 LysValaapSerLysAapAapLysGlyPheGlnA:argPheLeuAapThrSerGlnTyrLys 229  
 :::  
 DB 933 AAGGTCAGAGTCACAGAGAGCAAGGGCTTCAGAGATCCCTGGACAATGTCAGTACAAA 874  
 QY 230 CysAenSerLleLeuA:argTyrGluA:argValPheGlyProGlyTyrValSerThrGlyGly 249  
 :::  
 DB 873 TCCACTGGAAATCTTCGCTATAGCGGCTGTGGGAG-GGTTATGTAGACACCGGTGA 815  
 QY 250 TyrGluThrThrLysGluPheValSerMetLeuAapLysProGlyGlnLysValLeu 269  
 :::  
 DB 814 TTCGAGACCAAGAGGAGTTGTGGACACACTGGACTGAAGCTGTCAGAGGTGCTC 755  
 QY 270 AspValGlyCysGlyLleGlyGlyA:spPheTyrMetAlaGluThrPheAapValGlu 289  
 :::  
 DB 754 GATGTAGGGTGTGGTATCGGAGGAGCGGACTTCTACATGGCTGAAACCTATGATGTCCAT 695

QY 290 ValValGlyPheAapLeuSerValAenMetLleSerPheAlaLeuGluA:argSerLleGly 309  
 :::  
 DB 694 GTCTCGGAGATCGATCTTCTATCAACATGGTTCGTTGGTTCGATTTAGCGTGCATCGG 635  
 QY 310 LeuLysCyAlaValGluPheGluValAlaAapCyThrLysLleAenTyrProAapAen 329  
 :::  
 DB 634 CGCTCGTGTCTCGTGTGAGTTGAGTTCTGCTGACACCAAGAGGANTACGAGAGAAC 575  
 QY 330 SerPheAapValLleTyrSerArgAapThrLleLeuHisLleGlnAapLysProAlaLeu 349  
 :::  
 DB 574 ACCTTTGATGTCATCTACAGCCGTCAGACCATCTCCACATTCACAGACAAACCTGCTCTG 515  
 QY 350 PheArgSerPheTyrLysTyrLeuLysProGlyLysValLeuLleSerAapTyrCys 369  
 :::  
 DB 514 TTCAAGAACTTCTTCAGTGGCTCAGCGCTGGGGGCAAGTGTCTGATCATGACTACTGTC 455  
 QY 370 LysLysAlaGlyProProSerProGluPheAlaAlaTyrLleLysGlnA:argGlyTyrAsp 389  
 :::  
 DB 454 AGGAGCCCTGGGACACCATCAGAGGANTTCGCTGTATCATCAAGCAGAGAGGCTATGAC 395  
 QY 390 LeuHisAapValLysGluTyrGlyGlnMetLeuLysAapAlaGlyPheValAapValLeu 409  
 :::  
 DB 394 CTCATGATGTGAAGACCTACCGAAAGATGCTTGAGGATGCCGTTTCCATGATGTCGTC 335  
 QY 410 AlaGluAapArgThrGluGlnPheLleArgValLeuArgLysGluLeuGluThrValGlu 429  
 :::  
 DB 334 GCTGAGAGCCGACCCACCGACCATCTCTGAGGGTCTCTGAGAGGGAGCTGGGGAGACCGAG 275  
 QY 430 LysGluLysAapValPheLleSerAapPheSerGluGluAapTyrAenAapLleValGly 449  
 :::  
 DB 274 AAGAACAAAGAGGAGCTTCTCGCGGACTTACCCAGGAGGACTACGACGATCGTCAAC 215  
 QY 450 GlyTyrAenAapLysLeuA:argThrAlaLysGlyGluGlnA:argTyrGlyLeuPheVal 469  
 :::  
 DB 214 GCGCGAGCCGAGTTGAGCGGAGCTCTGCGCGAGAGCAGAGAGTGGGGCTTCTTCATC 155  
 QY 470 AlaLysLys 472  
 DB 154 GCGACCAAG 146

RESULT 8  
 CF513471 869 bp mRNA linear EST 09-SEP-2003  
 Locus Cabud0007 IF A02 Vitis vinifera cv. Cabernet sauvignon (Clone 8)  
 DEFINITION Bud - CABUD Vitis vinifera cDNA clone Cabud0007\_IF\_A02 5', mRNA  
 sequence.  
 ACCESSION CF513471  
 VERSION CF513471.1 GI:34545239  
 KEYWORDS EST.  
 SOURCE Vitis vinifera  
 ORGANISM Vitis vinifera  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; Vitaceae; Vitis.  
 REFERENCE 1 (bases 1 to 869)  
 AUTHORS Goes da Silva, P., Landolino, A., Lim, H., Baek, J., Jones, K. and  
 Cook, D.  
 TITLE Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon'  
 berries at various developmental stages  
 JOURNAL Unpublished (2003)  
 COMMENT Contact: Douglas Cook, PhD  
 CARS Genome Facility  
 UC Davis, Plant Pathology  
 One Shields Ave, Davis, CA 95616, USA  
 Tel: 530 754 6561  
 Fax: 530 754 6617  
 Email: drcook@ucdavis.edu  
 Seq primer: ACGGTACCGACATATGCC.  
 FEATURES ..... Location/Qualifiers  
 1. 869  
 source  
 /organism="Vitis vinifera"  
 /mol\_type="mRNA"











/lab\_host="SOLR"

/clone\_lib="mixed potato tissues"  
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
XhoI; supplier: Combination of untreated and phytophthora  
infestans-treated libraries of stolons, leaves, leaflets,  
axillary buds of stem explants, petioles, germinating  
eyes, tubers, or roots."

## ORIGIN

## Alignment Scores:

Pred. No.: 1.99e-121 Length: 706  
Score: 1057.00 Matches: 132  
Percent Similarity: 88.89% Conservative: 16  
Best Local Similarity: 82.05% Mismatches: 26  
Query Match: 42.03% Indels: 0  
DB: 13 Gaps: 0

US-10-031-331B-40 (1-473) x BQ120291 (1-706)

QY 156 ThrHisTyrArgGluProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAsp 175  
Db 4 ACCCATATATCGGAGCTTAGATTTTACAAAGGTTTAAAGATGTCATATAATGCT 63  
QY 176 GlySerGlyAenSerTyrGluLeuSerLeuSerCysLysCysAlleGlyAlaTyrVal 195  
Db 64 GGTGATGTAATCAATTGAACCTTCTCTCATCGTTGCAAGTCGATGGAGCTTATGTT 123  
QY 196 ArgAenLysAenGlnAenGlnLysSerTyrLeuTyrGlnLysValAspSerLysAsp 215  
Db 124 AAAAAAAGATATAGATATGATTTGTTGCTATGCAAGATGATTAATCTGCGAT 183  
QY 216 AspLysGlyPheGlnArgPheLeuAspThrSerGlnTyrLysCysAenSerLysLeuArg 235  
Db 184 GACAGGGGATTCGAGCGTTCTCGACATGTTTCAATACAAATGTAGTGCGCATCTGCGA 243  
QY 236 TyrGluArgValPheGlyProGlyTyrValSerThrGlyGlyTyrGlnThrLysGlu 255  
Db 244 TATGAACGTGCTTTGAGAGGTTATGTCAGCACAGGAGGCTTGATACACAAAGAA 303  
QY 256 PheValSerMetLeuAspLeuLysProGlyGlnLysValLeuAspValGlyCysGlyIle 275  
Db 304 TTGCTTTCTATGTTGATCTTCAACCTGCGCAAAAGTCTTGATGTTGGCTGTGGATA 363  
QY 276 GlyGlyAspPheTyrMetAlaGluThrPheAspValGluValGlyPheAspLeu 295  
Db 364 GGTGAGGTGACTTTTACATGCTGAGATTTATGATGTTTCAATGTTGCTTATTCACCTC 423  
QY 296 SerValAenMetLysPheAlaLeuGluArgSerLysGlyLysCysAlaValGlu 315  
Db 424 TCGATTACATGATATGTTGCTTCTGACGTTCAATGCTTCAATGTCCTGATG 483  
QY 316 PheGluValAlaAspCysThrLysLysLysLysLysLysLysLysLysLysLysLys 335  
Db 484 TTGAGGTGCTGATTTGATACAAAGAAACATATCTGATGCGCACATTTGATGTTTAC 543  
QY 336 SerArgAspThrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 355  
Db 544 AGTCGAGACATCTCTTCATCTCAACGACAAACCCACGTTTTCAGATCTTCTACAG 603  
QY 356 TrpLeuLysProGlyGlyValLeuLysSerAspTyrCysLysLysAlaGlyProPro 375  
Db 604 TGCTTGAACAGGAGGCAAGTCTCTAATGATGATTTACTGCAATATGCCAATACCA 663  
QY 376 SerProGluPheAlaAlaTyrLysLysLysLysLysLysLysLysLysLysLysLys 389  
Db 664 TCGGCAAGTTTCTGATATATAAAGCAAAAGGGGTTATGAT 705

## RESULT 14

BE040460

LOCUS

DEFINITION 'OE06G12' DE *Oryza sativa* cDNA 5' similar to methyltransferase, mRNA

sequence.

ACCESSION

BE040460

BE040460.1 GI:8335582

EST.

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

Source

1. 805

/organism="Oryza sativa"

/mol\_type="mRNA"

/strain="Pekali"

/db\_xref="taxon:4530"

/census\_type="roots"

/dev\_stage="1 week"

/clone\_lib="OR"

/notes="2-3 d 150mM NaCl"

Alignment Scores:

Pred. No.: 4.42e-121 Length: 805

Score: 1055.00 Matches: 198

Percent Similarity: 83.46% Conservative: 24

Best Local Similarity: 74.44% Mismatches: 44

Query Match: 41.95% Indels: 1

DB: 10 Gaps: 0

US-10-031-331B-40 (1-473) x BE040460 (1-805)

QY 160 GluProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAen 179

Db 3 GAGCCAGGTTCTATACAAAGATATTTAAAGATGCTTCTATGATATAAGATGGCGT 62

QY 180 SerTyrGluLeuSerLeuLeuSerCysLysCysAlleGlyAlaTyrValArgAenLysLys 199

Db 63 TCTTATGACCTTCTTAGAACATGCTGCTTGGGCTTATGTGAAAGCAGAA 122

QY 200 AenGlnAenGlnLysSerTyrLeuTyrGlnLysValAspSerLysAenAspLysGlyPhe 219

Db 123 AATCAATCATGTTATGTTGCTATGCGAAAGGTTTAACTCAACAGACAGACAGATTC 182

QY 220 GlnArgPheLeuAspThrSerGlnTyrLysCysAenSerLysLeuArgTyrGluArgVal 239

Db 193 CAAGATTTCTGACATGTCAGTACAAACCACTGGATCTTACGCTATGAGGCTGC 242

QY 240 PheGlyProGlyTyrValSerThrGlyGlyTyrGlnThrLysGluPheValSerMet 259

Db 243 TTGAGAGAGGTTTATGTCAGCATCTGTTGGAATTTGAACCAACCAAGGATTTTGTGATTAAG 302

QY 260 LeuAspLeuLysProGlyGlnLysValLeuAspValGlyCysGlyIleGlyGlyAsp 279

Db 303 CTGGATCTTAAACCTGGACAGAAAGTCTTGTATGTTGGTGGGATTTGAGAGCGCGCAC 362

QY 280 PheTyrMetAlaGluThrPheAspValGluValGlyPheAspLeuSerValAenMet 299

Db 363 TTTTATATGCTGGAAGTACGATGCTTCTGATGATGATGATGATGATGATGATGATG 422

QY 300 IleSerPheAlaLeuGluArgSerLysGlyLysCysAlaValGluPheGluValAla 319

Db 300 TTTTATATGCTGGAAGTACGATGCTTCTGATGATGATGATGATGATGATGATGATG 422



423 GTTTCATTGCAATCGAAGCGTCCATTGGACGCAAGTGTTCGGTTGAGTTGAAGTTGCT 482  
 QY AspCysThrLysIleAsnTyrProAspSerPheAspValIleTyrSerArgAspThr 339  
 DB GATTCACCAACCAAGACCTACGCCAACAATCATTTGATGTGATCTACAGCGGTGACACC 542  
 QY IleLeuHisIleGlnAspLysProAlaLeuPheArgSerPheTyrLysTyrPheLysPro 359  
 DB ATTCTTCACATACATGATAAACCCTGCTNTGTTTCAGAGTTTCTTCAAGTGGCTGAAACCT 602  
 QY GlyLysValLeuIleSerAspTyrCysLysLysAlaGlyProProSerProGluPhe 379  
 DB GNGGCAAGTCTCTCATCTAGTATCTAGCAATCTCGGAAACCATCAGAGAATTT 662  
 QY AlaAlaTyrLysGlnArgGlyTyrAspLeuHisAspValLysGluTyrGlyGlnMet 399  
 DB GCTCTTACATTAAGCGAGAGGCTATGACCTCCAGATGTGAGACTTACCGAGAGATG 722  
 QY LeuLysAspAlaGlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIleArg 419  
 DB CTTGAGGATCTGGGGTTTCATCGCATTTGCTGAAGACGC-ACGACCAAGTCTCTCGCT 781  
 QY 420 ValLeuArgLysGluLeu 425  
 DB GTTCTTCAGAGGAGCTTG 799  
 RESULT 15  
 CF482195  
 LOCUS  
 DEFINITION  
 POL1\_5\_B02\_g1\_A002 Pollen Sorghum bicolor cDNA clone  
 POL1\_5\_B02\_A002 5', mRNA sequence.  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Sorghum bicolor (sorghum)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD  
 clade; Panicoideae; Andropogoneae; Sorghum.  
 1 (bases 1 to 721)  
 Cordonnier-Pratt, M.-M., Suzuki, Y., Sugano, S., Klein, R.R., Liang, C.,  
 Sun, F., Sullivan, R., Eastman, A., Cannon, R., Kern, B., Morgan, J.,  
 Lucas, A., Al-Sheikh, A., Jones, V., Adibi, N., Owen, A., Gao, J. and  
 Pratt, L.H.  
 EST database from Sorghum: pollen  
 Unpublished (2003)  
 Other\_ESTR: POL1\_5\_B02.b1\_A002  
 Contact: Cordonnier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmpratt@uga.edu  
 Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in  
 the Human Genome Center, University of Tokyo Institute of Medical  
 Science; plant material and RNA prepared at Texas A & M University;  
 sequencing done in the Laboratory for Genomics and Bioinformatics,  
 University of Georgia. Sequence ends have been trimmed to exclude  
 vector and regions below phred quality 16. Three-prime sequences  
 are presented as their reverse complement and have been trimmed to  
 exclude polyA.  
 Seq primer: Sug5 (CTTCGTCTCTAAAGCTGGC).  
 Location/Qualifiers  
 1..721  
 /organism="Sorghum bicolor"  
 /mol\_type="mRNA"  
 /cultivar="BTx623"  
 /db\_xref="taxon:4558"  
 /cclone="POL1\_5\_B02\_A002"  
 /lab\_host="DH10B-T1 phage-resistant E. coli"  
 /clone\_lib="Pollen"  
 /note="Organ: Pollen; Vector: pME18S-FLJ3; Site\_1: XhoI;"

Site 2: XhoI; The library was prepared from polyA+ RNA from pollen at the late vacuolated-vacuolated stage of development. Pollen was harvested from greenhouse-grown panicles of sorghum line BTx623. Panicles were removed from the flag leaf prior to emergence, when no detectable amylose is present in pollen of male-fertile lines. This stage represents pollen collected from anthers about 8-14 days prior to anthesis. Double-stranded cDNA was cloned unidirectionally into different DraIII sites of the pME18S-FLJ3 vector (5-prime DraIII site is CACTGTGTG, 3-prime DraIII site is CACCATGTG). XhoI excises the cDNA insert."

## ORIGIN

Alignment Scores:  
 Pred. No.: 5,02e-119 Length: 721  
 Score: 1038.00 Matches: 190  
 Percent Similarity: 89.92% Conservatives: 24  
 Best Local Similarity: 79.83% Mismatches: 24  
 Query Match: 41.27% Indels: 0  
 DB: 14 Gaps: 0

US-10-031-331B-40 (1-473) x CF482195 (1-721)

QY 235 ArgTyrGluArgValPheGlyProGlyTyrValSerThrGlyGlyTyrGluThrThrLys 254  
 DB 8 CGTTATGAGCGGTGTCTTTGGTGAAGGTTATGTGAGCACTGTGTGGTAATCGAGACTACAAAG 67

QY 255 GluPheValSerMetLeuAspLeuLysProGlyGlnLysValLeuAspValGlyCysGly 274  
 DB 68 GAATTTGTGGCATGCTGGATCTTAAACCTGGCAAGAAAGTACTTGATGTGTGGATGGGA 127

QY 275 IleGlyGlyGlyAspPheTyrMetAlaGluThrPheAspValGluValValGlyPheAsp 294  
 DB 128 ATTGAGCGCGGCACTTTACATGCTGNAACATGATGTCATGTCCTGCGATTGAT 187

QY 295 LeuSerValAsnMetIleSerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaVal 314  
 DB 188 CTTTGGGTTTAAACATGTTTCATTTGCAATGAACGTCGTCATTTGGACGCAAGTCTCTGTT 247

QY 315 GluPheGluValAlaAspCysThrLysIleAsnTyrProAspAsnSerPheAspValIle 334  
 DB 248 CAATTTGAGTTTCTGATGTCACCAACAAAGGATTATCCCAAGAGATAGTTTTCAGCTCATC 307

QY 335 TyrSerArgAspThrIleLeuHisIleGlnAspLysPheAlaLeuPheAspSerPheTyr 354  
 DB 308 TACAGCGGTGACACTATCTTTCATACAGACAGACGCTGCTCTGTTTTCAGAGCTTCTTC 367

QY 355 LysTrpLeuLysProGlyGlyLysValLeuIleSerAspTyrCysLysLysAlaGlyPro 374  
 DB 368 AAATGGCTTAAACCCCGGTGGCAAGTCTTAATCAGTGACTACTGTAGAAATCTCTGGAANA 427

QY 375 ProSerProGluPheAlaAlaTyrIleLysGlnArgGlyTyrAspLeuHisAspValLys 394  
 DB 428 CCNTCAGAGAAATTTGCTGCATATACATACAGAGAGGTTATGACCTCCACGATGTGAAG 487

QY 395 GluTyrGlyGlnMetLeuLysAspAlaGlyPheValAspValLeuAlaGluAspArgThr 414  
 DB 488 GCTTATGACAGATGCTCAAGGATGCTGTTTTCATGATGTCTTTCATGATGTCTGCAAGATGCGACT 547

QY 415 GluGlnPheIleArgValLeuArgLysGluLeuGluThrValGluLysGluLysAspVal 434  
 DB 548 GAGCAGTTCTGTAATGTTCTACCGAGGAGGAGTAGGTGAAGTTGNAAGACAAAGAGGCT 607

QY 435 PheIleSerAspPheSerGluGluAspTyrAsnAspIleValGlyGlyTyrPheAspLys 454  
 DB 608 TTCTTGGCAGACTTTCACCCAGGAGGACTACGACGACATTTGTAACGCTTGAACGCGGAAG 667

QY 455 LeuArgArgThrAlaLysGlyGluGlnArgTrpGlyLeuPheValAlaLysLys 472  
 DB 668 CTGAACCGGAGCTCTGCGCGGTGACGAGAGGTGGCGCTCTGTTTCATTGGACCAAG 721

Tue Aug 3 10:37:38 2004

us-10-031-331b-40.rst

Page 14

Search completed: August 2, 2004, 03:21:23  
Job time : 3593 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2004, 13:15:57 ; Search time 54 Seconds  
(without alignments)  
2474.907 Million cell updates/sec

Title: US-10-031-331b-40

Perfect score: 473

Sequence: 1 HTVDTLTIAMMLDSQASDL.....KLRTAKGEQRWGLFVAKKK 473

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1586107 seqs, 282547505 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A\_Geneseq\_29Jan04.\*  
1: Geneseqp1980s.\*  
2: Geneseqp1990s.\*  
3: Geneseqp2000s.\*  
4: Geneseqp2001s.\*  
5: Geneseqp2002s.\*  
6: Geneseqp2003as.\*  
7: Geneseqp2003bs.\*  
8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	473	100.0	473	4	AAB80627
2	45	9.5	289	4	AAE09761
3	45	9.5	494	4	AAE09760
4	23	4.9	191	7	ABM74064
5	21	4.4	168	4	ABM74064
6	21	4.4	491	4	ABM74064
7	16	3.4	287	7	ABM74064
8	16	3.4	346	5	AAE13618
9	14	3.0	293	3	AAE13618
10	14	3.0	293	3	AAE13618
11	14	3.0	293	3	AAE13618
12	14	3.0	293	3	AAE13618
13	14	3.0	314	5	ABM74064
14	14	3.0	336	3	AAE13618
15	14	3.0	336	3	AAE13618
16	14	3.0	336	3	AAE13618
17	14	3.0	339	3	AAE13618
18	14	3.0	358	3	AAE13618
19	14	3.0	361	3	AAE13618
20	14	3.0	365	3	AAE13618
21	14	3.0	366	3	AAE13618
22	11	2.3	137	3	AAE13618
23	11	2.3	140	3	AAE13618
24	11	2.3	152	3	AAE13618
25	11	2.3	202	6	ABO01277

26	11	2.3	344	2	AAW70838
27	11	2.3	344	3	AAW80992
28	10	2.1	301	4	AAW46833
29	10	2.1	309	6	ABP96400
30	10	2.1	318	3	AAW69405
31	10	2.1	318	4	AAW69405
32	10	2.1	318	4	AAW69405
33	10	2.1	318	4	AAW69405
34	10	2.1	318	5	AAW69405
35	9	1.9	280	6	ABP96396
36	9	1.9	280	6	ABP96396
37	9	1.9	280	7	ADD19157
38	9	1.9	280	7	ADD19157
39	9	1.9	311	6	ABP96401
40	9	1.9	330	6	ABP96399
41	9	1.9	333	4	ABG30300
42	9	1.9	621	4	ABG30299
43	9	1.9	653	4	ABG18352
44	8	1.7	43	4	AAE82902
45	8	1.7	43	6	ABR42164

## ALIGNMENTS

## RESULT 1

AAAB80627  
ID AAB80627 standard; protein; 473 AA.

XX AAB80627;

DT 06-AUG-2003 (revised)

DT 02-MAY-2001 (first entry)

XX Environmental stress tolerant protein SEQ ID 40.

XX Environmental stress resistance; salt; heat; desert; transgenic plant.

XX Suaeda japonica.

XX WO200106006-A1.

XX 25-JAN-2001.

XX 19-JUL-2000; 2000WO-JP004862.

XX 19-JUL-1999; 99JP-00235910.

XX 24-MAR-2000; 2000JP-00085377.

XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.

XX Yamada A, Ozeki Y, Saito T;

XX WPI; 2001-147355/15.

XX N-PSDB; AAF74206.

Screening method to obtain DNA encoding environmental stress resistance factor, useful for producing transgenic plants resistant to environmental stress.

Claim 64; Page 125-127; 167pp; Japanese.

Polynucleotide sequences AAF74187 - AAF74218 encode proteins AAB80608 - AAB80639, which impart environmental stress resistance. The invention relates to a method for identifying DNA encoding proteins imparting environmental stress resistance. The method comprises inserting cDNA from a library originating in a salt-resistant organism into a host cell, culturing the transformants under conditions in which the untransformed host does not grow well, and selecting for viable clones. The method is useful for obtaining DNA encoding environmental stress resistance factors. The DNA encoding proteins conferring environmental stress resistance, can be used in the production of plants resistant to environmental stress, which can be cultivated in unfavourable

CC environments such as deserts, salt damaged ground, cold regions and the  
 CC oceans. They can be used for increasing the area of land covered by green  
 CC plants, and desert greening and afforestation, in order to counter the  
 CC effects of the increase in atmospheric carbon dioxide concentration. PCR  
 CC primers AAF74219 and AAF74220 are used in an example illustrating the  
 CC method of the invention. (Updated on 06-AUG-2003 to correct OS field.)  
 XX

XX Sequence 473 AA;

Query Match 100.0%; Score 473; DB 4; Length 473;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 473; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HTVLTTEAMMLDSQASDLDEKPERPILSMLEKCKLLEAGIGRFTGLAEKAGQV 60  
 DB 1 HTVLTTEAMMLDSQASDLDEKPERPILSMLEKCKLLEAGIGRFTGLAEKAGQV 60  
 QY 61 IALDPTESAIKONEVINGHYKVNKPMCAVDVTSTLSPPHSLDVI FSNWLLMYLSDRVE 120  
 DB 61 IALDPTESAIKONEVINGHYKVNKPMCAVDVTSTLSPPHSLDVI FSNWLLMYLSDRVE 120  
 QY 121 NLVERMLKWLKPGGYIFFRESCFHQSDHKKSNPTHYREPRFTYKAFKECHLQDGSNS 180  
 DB 121 NLVERMLKWLKPGGYIFFRESCFHQSDHKKSNPTHYREPRFTYKAFKECHLQDGSNS 180  
 QY 181 YELSLSCCKICAVRNKQNOISLWQKVDSDKDFORPLDTSQYKCNLSLYRVP 240  
 DB 181 YELSLSCCKICAVRNKQNOISLWQKVDSDKDFORPLDTSQYKCNLSLYRVP 240  
 QY 241 GGVYSTGYYETTKFVSMLEKPGQKVLVDGCGIGGDFYMAETFDVEVWGPOLSVNMI 300  
 DB 241 GGVYSTGYYETTKFVSMLEKPGQKVLVDGCGIGGDFYMAETFDVEVWGPOLSVNMI 300  
 QY 301 SFALERSIGLKCAVEFEVADCTKINYPDNSFDVYSRDTILHI QKPALEFSFYKWLKPG 360  
 DB 301 SFALERSIGLKCAVEFEVADCTKINYPDNSFDVYSRDTILHI QKPALEFSFYKWLKPG 360  
 QY 361 GKVLSIDYCKKAGPPSPPEPAAIYKQGYDLHDVWYQMLKQAGFVDVLAEDRTEQFIRV 420  
 DB 361 GKVLSIDYCKKAGPPSPPEPAAIYKQGYDLHDVWYQMLKQAGFVDVLAEDRTEQFIRV 420  
 QY 421 LRKELETVEKDVIFISDFSEEDYNDIVGWMNDKLARTANGQRNGLFVAKKK 473  
 DB 421 LRKELETVEKDVIFISDFSEEDYNDIVGWMNDKLARTANGQRNGLFVAKKK 473

RESULT 2

AAE09761  
 ID AAE09761 standard; protein; 289 AA.

XX AAE09761;

XX 29-NOV-2001 (first entry)

XX Spinach PEAMT truncated protein.

XX Spinach; S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase;  
 KW PEAMT; cellular intermediate; phospho-dimethylethanolamine; choline;  
 KW phosphomono-methylethanolamine; phosphocholine; phosphatidylcholine;  
 KW glycine betaine; choline-O-sulphate; lipid content alteration;  
 KW osmotic stress tolerance; nutritional value; transgenic plant;  
 KW cryoprotectant.

XX Spinacia oleracea.

XX WO200168870-A2.

XX 20-SEP-2001.

XX 15-MAR-2001; 2001WO-US008352...

XX 15-MAR-2000; 2000US-00525885.

XX

PA (UYFL) UNIV FLORIDA.  
 PA (UYCA-) UNIV CARNEGIE MELLON.

PI Hanson AD, Nuccio ML, Henry SA;

XX WPI; 2001-565796/63.

DR N-PSDB; AAD16798.

XX New plant S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase  
 PT polypeptides, useful for modulating the levels of cellular intermediates  
 PT such as phosphodimethylethanolamine and for altering the lipid content in  
 PT plants cells.

XX Claim 1; Page 109; 158pp; English.

XX The present sequence is S-adenosyl-L-methionine:phosphoethanolamine N-  
 CC methyltransferase (PEAMT) truncated protein from spinach. The PEAMT  
 CC sequences are useful for modulating the levels of cellular intermediates  
 CC such as phosphodimethylethanolamine, phosphomono-methylethanolamine,  
 CC choline, phosphocholine, phosphatidylcholine, choline-O-sulphate or  
 CC glycine betaine. They are useful for altering the lipid content in plant  
 CC cells. The polynucleotides are also useful for improving the osmotic  
 CC stress tolerance of a plant and increasing the cryoprotectant properties  
 CC of a plant. The present invention also relates to methods and  
 CC compositions comprising PEAMT used for generating transgenic plants with  
 CC increased nutritional value

XX Sequence 289 AA;

Query Match 9.5%; Score 45; DB 4; Length 289;  
 Best Local Similarity 100.0%; Pred. No. 8.1e-40;  
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 LVERMLKWLKPGGYIFFRESCFHQSDHKKSNPTHYREPRFTYK 166  
 DB 143 LVERMLKWLKPGGYIFFRESCFHQSDHKKSNPTHYREPRFTYK 187

RESULT 3

AAE09760

ID AAE09760 standard; protein; 494 AA.

XX AAE09760;

XX 29-NOV-2001 (first entry)

XX Spinach PEAMT protein.

XX Spinach; S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase;  
 KW PEAMT; cellular intermediate; phospho-dimethylethanolamine; choline;  
 KW phosphomono-methylethanolamine; phosphocholine; phosphatidylcholine;  
 KW glycine betaine; choline-O-sulphate; lipid content alteration;  
 KW osmotic stress tolerance; nutritional value; transgenic plant;  
 KW cryoprotectant.

XX Spinacia oleracea.

XX Key Location/Qualifiers

FT Misc-difference 462

FT /note= "This residue is given as Lys in the sequence  
 FT shown as SEQ ID NO: 2 in figure 3 of the specification"

XX WO200168870-A2.

XX 20-SEP-2001.

XX 15-MAR-2001; 2001WO-US008352.

XX 15-MAR-2000; 2000US-00525885.

XX (UYFL) UNIV FLORIDA.

PA (UYCA-) UNIV CARNEGIE MELLON.

XX

PI Hanson AD, Nuccio ML, Henry SA;  
 XX WPI; 2001-565796/63.  
 DR N-PSDB; RAD16797.  
 XX  
 XX New plant S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase  
 PT polypeptides, useful for modulating the levels of cellular intermediates  
 PT such as phosphodimethylmethanolamine and for altering the lipid content in  
 PT plant cells.  
 XX  
 PS Claim 1; Page 109; 158pp; English.  
 XX  
 CC The present sequence is S-adenosyl-L-methionine:phosphoethanolamine N-  
 CC methyltransferase (PEMT) protein from spinach. The PEMT sequences are  
 CC useful for modulating the levels of cellular intermediates such as  
 CC phosphodimethylmethanolamine, phosphomono-methylmethanolamine, choline,  
 CC phosphocholine, phosphatidylcholine, choline-O-sulphate or glycine  
 CC betaine. They are useful for altering the lipid content in plant cells.  
 CC The polynucleotides are also useful for improving the osmotic stress  
 CC tolerance of a plant and increasing the cryoprotectant properties of a  
 CC plant. The present invention also relates to methods and compositions  
 CC comprising PEMT used for generating transgenic plants with increased  
 CC nutritional value  
 XX  
 SQ Sequence 494 AA;  
 Query Match 9.5%; Score 45; DB 4; Length 494;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-39;  
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 122 LVERMLKPGGVIFFRESCFQSGDHKSNPTHYRPRFTK 166  
 DB 143 LVERMLKPGGVIFFRESCFQSGDHKSNPTHYRPRFTK 187  
 RESULT 4  
 ABM74064  
 ID ABM74064 standard; protein; 191 AA.  
 AC ABM74064;  
 XX  
 XX 17-OCT-2003 (first entry)  
 DE DNA clone originating in barley containing SNP sequence #474.  
 XX  
 XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis.  
 XX  
 XX Hordeum vulgare.  
 XX  
 XX WO2003057877-A1.  
 XX  
 XX 17-JUL-2003.  
 XX  
 XX 16-DEC-2002; 2002WO-IB005403.  
 XX  
 XX 20-DEC-2001; 2001JP-00387059.  
 XX  
 XX 20-DEC-2001; 2001JP-00387131.  
 XX  
 XX 20-DEC-2001; 2001JP-00403299.  
 XX  
 XX 20-DEC-2001; 2001JP-00403300.  
 XX  
 XX 27-SEP-2002; 2002JP-00327515.  
 XX  
 XX (UNIV-) UNIV JAPAN OKAYAMA.  
 XX  
 XX Sato K, Takeda K, Kohara Y;  
 PT WPI; 2003-587127/55.  
 XX  
 XX Single nucleotide polymorphism sites in barley varieties and DNA  
 PT sequences containing them for analysis and identification of barley  
 PT varieties and production of barley transformants with desired  
 PT characteristics.  
 XX  
 PS Disclosure; SEQ ID XX; 284pp; Japanese.

XX The present invention relates to oligonucleotide clones originating in  
 CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms  
 CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley  
 CC varieties, identification of particular varieties and genotype-phenotype  
 CC analysis, isolation of specific genes and creation of new varieties by  
 CC transformation of barley varieties with them and production of new barley  
 CC varieties with desired properties. The present sequence represents an  
 CC oligonucleotide clone sequence featured in the specification. The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published-pct-sequences  
 XX  
 SQ Sequence 191 AA;  
 Query Match 4.9%; Score 23; DB 7; Length 191;  
 Best Local Similarity 100.0%; Pred. No. 5.4e-16;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 331 FDVIYSRDTLHLDKPKALFRSP 353  
 DB 50 FDVIYSRDTLHLDKPKALFRSP 72  
 RESULT 5  
 AAB99873  
 ID AAB99873 standard; protein; 168 AA.  
 XX  
 XX AAB99873;  
 XX  
 XX 21-SEP-2001 (first entry)  
 DT  
 XX  
 DE Physcomitrella patens 78\_pppprotl\_092\_e12rev protein.  
 XX  
 XX Tocopherol and carotenoid metabolism related protein; TCWRP; synthesis;  
 XX Physcomitrella patens; moss; algae; microorganism; fungus; plant;  
 XX identification; genome mapping; modulation; evolutionary study;  
 XX cellular production; fine chemical.  
 XX  
 OS Physcomitrella patens.  
 XX  
 XX WO200144276-A2.  
 XX  
 XX 21-JUN-2001.  
 XX  
 XX 14-DEC-2000; 2000WO-EP012698.  
 XX  
 XX 16-DEC-1999; 99US-0171121P.  
 XX  
 XX (BADI) BASF PLANT SCI GMBH.  
 XX  
 XX Lerchl J, Renz A, Ehrhardt T, Reindl A, Cirpus P, Blaschoff F;  
 XX Frank M, Freund A, Duwenig E, Schmidt R, Reski R, Badur R;  
 DR WPI; 2001-398121/42.  
 DR N-PSDB; AAH44246.  
 XX  
 XX Tocopherol and carotenoid metabolism related protein (TCWRP), used to  
 PT produce fine chemicals, is isolated from mosses, algae, microorganisms,  
 PT fungi, plants, or their fragments.  
 XX  
 PS Claim 28; Page 118-119; 123pp; English.  
 XX  
 CC The present invention describes isolated tocopherol and carotenoid  
 CC metabolism related proteins (TCWRP) (I) from mosses or algae,  
 CC microorganisms or fungi, plants, or its fragments. (I) can be used as  
 CC enzymes in the production of fine chemicals or in the metabolism of  
 CC tocopherols and carotenoids. (I) also assist in transmembrane transport.  
 CC The fine chemicals that can be produced include lipids, fatty acids,  
 CC vitamins, cofactors, enzymes, amino acids, and nucleotide bases.  
 CC Nucleotide sequences, proteins, vectors and host cells from the present  
 CC invention can be used: (a) to identify mosses related to Physcomitrella  
 CC patens; (b) in mapping genomes of mosses related to Physcomitrella patens

CC ; (c) in the modulation of TCMP activity; (d) in evolutionary studies;  
 CC (e) in the determination of functional TCMP regions; (f) and in the  
 CC cellular production of fine chemicals. AAH44222 to AAH44262 encode the  
 CC Physcomitrella patens TCMP proteins given in AAB99889 to AAB99889.  
 CC AAH44212 to AAH44221 represent nucleotide sequence used in the  
 CC exemplification of the present invention

XX SQ Sequence 168 AA;  
 Query Match 4.4%; Score 21; DB 4; Length 168;  
 Best Local Similarity 100.0%; Pred. No. 7.3e-14;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 330 SFQVYSKDTILHIDKPALE 350  
 DB 24 SFQVYSKDTILHIDKPALE 44

RESULT 6  
 AAB99889  
 ID AAB99889 standard; protein; 491 AA.  
 XX AAB99889;  
 XX 21-SEP-2001 (first entry)  
 XX Physcomitrella patens 78\_ppprot1\_092\_e12-260rev protein.

XX Tocopherol and carotenoid metabolism related protein; TCMP; synthesis;  
 XX Physcomitrella patens; moss; algae; microorganism; fungus; plant;  
 XX identification; genome mapping; modulation; evolutionary study;  
 XX cellular production; fine chemical.

XX Physcomitrella patens.  
 XX WO200144276-A2.  
 XX 21-JUN-2001.

XX 14-DEC-2000; 2000WO-EP012698.  
 XX 16-DEC-1999; 99US-0171121P.  
 XX (BADI ) BASF PLANT SCI GMBH.

XX Lerchl J, Renz A, Ehrhardt T, Reindl A, Cirpus P, Bischoff F;  
 XX Frank M, Freund A, Duwenig E, Schmidt R, Reski R, Badur R;  
 XX WPI; 2001-398121/42.  
 XX N-PSDB; AAH44262.

XX Tocopherol and carotenoid metabolism related protein (TCMP), used to  
 XX produce fine chemicals, is isolated from mosses, algae, microorganisms,  
 XX fungi, plants, or their fragments.

XX Claim 28; Page 122-123; 123pp; English.

XX The present invention describes isolated tocopherol and carotenoid  
 XX metabolism related proteins (TCMP) (I) from mosses or algae,  
 XX microorganisms or fungi, plants, or its fragments. (I) can be used as  
 XX enzymes in the production of fine chemicals or in the metabolism of  
 XX tocopherols and carotenoids. (I) also assist in transmembrane transport.  
 XX The fine chemicals that can be produced include lipids, fatty acids,  
 XX vitamins, cofactors, enzymes, amino acids, and nucleotide bases.  
 XX Nucleotide sequences, proteins, vectors and host cells from the present  
 XX invention can be used; (a) to identify mosses related to Physcomitrella  
 XX patens; (b) in mapping genomes of mosses related to Physcomitrella patens  
 XX ; (c) in the modulation of TCMP activity; (d) in evolutionary studies;  
 XX (e) in the determination of functional TCMP regions; (f) and in the  
 XX cellular production of fine chemicals. AAH44222 to AAH44262 encode the  
 XX Physcomitrella patens TCMP proteins given in AAB99889 to AAB99889.  
 XX AAH44212 to AAH44221 represent nucleotide sequence used in the  
 XX exemplification of the present invention

XX SQ Sequence 491 AA;  
 Query Match 4.4%; Score 21; DB 4; Length 491;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-13;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 PGGYIPFRSCFHSGDHKKK 152  
 DB 148 PGGYIPFRSCFHSGDHKKK 168

RESULT 7  
 ABM74462  
 ID ABM74462 standard; protein; 287 AA.  
 XX ABM74462;  
 XX 17-OCT-2003 (first entry)  
 XX DNA clone originating in barley containing SNP sequence #872.  
 XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis.  
 XX Hordeum vulgare.  
 XX WO2003057877-A1.  
 XX 17-JUL-2003.  
 XX 16-DEC-2002; 2002MO-IB005403.  
 XX 20-DEC-2001; 2001JP-00387059.  
 XX 20-DEC-2001; 2001JP-00387131.  
 XX 20-DEC-2001; 2001JP-00403299.  
 XX 20-DEC-2001; 2001JP-00403300.  
 XX 27-SEP-2002; 2002JP-00327315.  
 XX (UTMI-) UNIV JAPAN OKAYAMA.  
 XX Sato K, Takeda K, Kohara Y;  
 XX WPI; 2003-587127/55.  
 XX Single nucleotide polymorphism sites in barley varieties and DNA  
 XX sequences containing them for analysis and identification of barley  
 XX varieties and production of barley transformants with desired  
 XX characteristics.

XX Disclosure; SEQ ID XX; 284pp; Japanese.

XX The present invention relates to oligonucleotide clones originating in  
 XX barley (Hordeum vulgare) which contain single nucleotide polymorphisms  
 XX (SNP). The oligonucleotides may be used for analysis of SNPs among barley  
 XX varieties, identification of particular varieties and genotype-phenotype  
 XX analysis, isolation of specific genes and creation of new varieties by  
 XX transformation of barley varieties with them and production of new barley  
 XX varieties with desired properties. The present sequence represents an  
 XX oligonucleotide clone sequence featured in the specification. The  
 XX sequence data for this patent did not form part of the printed  
 XX specification, but was obtained in electronic format directly from WIPO  
 XX at ftp.wipo.int/pub/published-pct-sequences

XX SQ Sequence 287 AA;  
 Query Match 3.4%; Score 16; DB 7; Length 287;  
 Best Local Similarity 100.0%; Pred. No. 3.5e-08;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 IFSNWLMLYLSDEEVE 120  
 DB 141 IFSNWLMLYLSDEEVE 156

```

RESULT 8
AAE13618
ID AAE13618 standard; protein; 346 AA.
AC AAE13618;
XX
XX
XX 07-MAR-2002 (first entry)
XX Nicotiana tabacum smt1 (Ntsmt1) full length protein.
XX Nicotiana tabacum smt1 (Ntsmt1) full length protein.
XX Sterol methyl transferase 1; SMT1 gene; cholesterol; oilseed; tobacco.
XX Nicotiana tabacum.
XX WO200179513-A2.
XX
XX PD 25-OCT-2001.
XX
XX 04-APR-2001; 2001WO-EP003808.
XX 14-APR-2000; 2000EP-00303193.
XX (UNIL ) UNILEVER NV.
XX (UNIL ) UNILEVER PLC.
XX (UNIL ) HINDUSTAN LEVER LTD.
XX
XX PI Harker M, Hellyer SA; Holmber N, Safford D;
XX
XX WPI: 2002-034363/04.
XX N-PSDB; AAD23672.
XX
XX Using a gene expressing sterol methyl transferase 1 to increase the level
XX of sterols in plant seeds or decrease cholesterol level in plant tissue
XX PT is useful to produce commercially desired plants.
XX
XX Example 1; Fig 1B; 50pp; English.
XX
XX The invention relates to the use of a gene expressing sterol methyl
XX transferase 1 (SMT1) to increase the level of sterols in plant seeds or
XX decrease cholesterol level in plant tissue. SMT1 genes of the invention
XX are used to produce commercially desired plants and oilseeds where the
XX plant tissue contains reduced cholesterol and the oilseeds contain
XX sterols that may be extracted along with the oils. The present sequence
XX is Nicotiana tabacum smt1 (Ntsmt1) full length protein
XX
XX Sequence 346 AA;
XX
XX Query Match 3.4%; Score 16; DB 5; Length 346;
XX Best Local Similarity 100.0%; Pred. No. 4.2e-08;
XX Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 262 LKPGQKVLVDVGGIGG 277
XX Db 98 LKPGQKVLVDVGGIGG 113
XX
XX
XX RESULT 9
XX AAG45955
XX ID AAG45955 standard; protein; 293 AA.
XX AC AAG45955;
XX
XX 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 57758.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; Genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX

```

```

PN EP1033405-A2.
XX
XX PD 06-SEP-2000.
XX
XX PF 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 28-APR-1999; 99US-0130891P.
XX 30-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 04-MAY-1999; 99US-0132484P.
XX 05-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 07-MAY-1999; 99US-0132487P.
XX 11-MAY-1999; 99US-0132863P.
XX 14-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138034P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138847P.
XX 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 16-JUN-1999; 99US-0139453P.
XX 17-JUN-1999; 99US-0139492P.
XX 18-JUN-1999; 99US-0139454P.
XX 18-JUN-1999; 99US-0139455P.
XX 18-JUN-1999; 99US-0139456P.
XX 18-JUN-1999; 99US-0139457P.
XX 18-JUN-1999; 99US-0139458P.
XX 18-JUN-1999; 99US-0139459P.
XX 18-JUN-1999; 99US-0139460P.
XX 18-JUN-1999; 99US-0139461P.
XX 18-JUN-1999; 99US-0139462P.
XX 18-JUN-1999; 99US-0139463P.
XX 18-JUN-1999; 99US-0139750P.
XX 18-JUN-1999; 99US-0139763P.
XX 21-JUN-1999; 99US-0139817P.
XX 22-JUN-1999; 99US-0139899P.
XX 23-JUN-1999; 99US-0140353P.
XX 23-JUN-1999; 99US-0140354P.
XX 24-JUN-1999; 99US-0140695P.
XX 28-JUN-1999; 99US-0140823P.
XX 28-JUN-1999; 99US-0140891P.
XX 29-JUN-1999; 99US-0141287P.
XX 01-JUL-1999; 99US-0141842P.
XX

```

PR 01-JUL-1999; 99US-0142154P.  
 PR 02-JUL-1999; 99US-0142055P.  
 PR 06-JUL-1999; 99US-0142300P.  
 PR 08-JUL-1999; 99US-0142803P.  
 PR 09-JUL-1999; 99US-0142920P.  
 PR 12-JUL-1999; 99US-0142977P.  
 PR 13-JUL-1999; 99US-0143542P.  
 PR 14-JUL-1999; 99US-0143624P.  
 PR 15-JUL-1999; 99US-0144005P.  
 PR 16-JUL-1999; 99US-0144085P.  
 PR 16-JUL-1999; 99US-0144086P.  
 PR 19-JUL-1999; 99US-0144325P.  
 PR 19-JUL-1999; 99US-0144331P.  
 PR 19-JUL-1999; 99US-0144332P.  
 PR 19-JUL-1999; 99US-0144333P.  
 PR 19-JUL-1999; 99US-0144334P.  
 PR 19-JUL-1999; 99US-0144335P.  
 PR 20-JUL-1999; 99US-0144352P.  
 PR 20-JUL-1999; 99US-0144353P.  
 PR 20-JUL-1999; 99US-0144354P.  
 PR 20-JUL-1999; 99US-0144884P.  
 PR 21-JUL-1999; 99US-0144814P.  
 PR 21-JUL-1999; 99US-0145085P.  
 PR 21-JUL-1999; 99US-0145086P.  
 PR 22-JUL-1999; 99US-0145087P.  
 PR 22-JUL-1999; 99US-0145088P.  
 PR 22-JUL-1999; 99US-0145089P.  
 PR 22-JUL-1999; 99US-0145090P.  
 PR 22-JUL-1999; 99US-0145129P.  
 PR 23-JUL-1999; 99US-0145145P.  
 PR 23-JUL-1999; 99US-0145218P.  
 PR 23-JUL-1999; 99US-0145224P.  
 PR 26-JUL-1999; 99US-0145276P.  
 PR 27-JUL-1999; 99US-0145913P.  
 PR 27-JUL-1999; 99US-0145914P.  
 PR 27-JUL-1999; 99US-0145915P.  
 PR 28-JUL-1999; 99US-0145916P.  
 PR 02-AUG-1999; 99US-0145951P.  
 PR 02-AUG-1999; 99US-0146386P.  
 PR 02-AUG-1999; 99US-0146387P.  
 PR 02-AUG-1999; 99US-0146388P.  
 PR 03-AUG-1999; 99US-0146389P.  
 PR 03-AUG-1999; 99US-0147038P.  
 PR 04-AUG-1999; 99US-0147204P.  
 PR 04-AUG-1999; 99US-0147205P.  
 PR 05-AUG-1999; 99US-0147192P.  
 PR 05-AUG-1999; 99US-0147260P.  
 PR 06-AUG-1999; 99US-0147303P.  
 PR 06-AUG-1999; 99US-0147316P.  
 PR 09-AUG-1999; 99US-0147493P.  
 PR 09-AUG-1999; 99US-0147935P.  
 PR 10-AUG-1999; 99US-0148171P.  
 PR 11-AUG-1999; 99US-0148319P.  
 PR 12-AUG-1999; 99US-0148341P.  
 PR 13-AUG-1999; 99US-0148565P.  
 PR 13-AUG-1999; 99US-0148684P.  
 PR 16-AUG-1999; 99US-0149368P.  
 PR 17-AUG-1999; 99US-0149175P.  
 PR 18-AUG-1999; 99US-0149426P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 20-AUG-1999; 99US-0149929P.  
 PR 23-AUG-1999; 99US-0149902P.  
 PR 23-AUG-1999; 99US-0149930P.  
 PR 23-AUG-1999; 99US-0150566P.  
 PR 26-AUG-1999; 99US-0150884P.  
 PR 27-AUG-1999; 99US-0151065P.  
 PR 27-AUG-1999; 99US-0151066P.  
 PR 27-AUG-1999; 99US-0151080P.  
 PR 30-AUG-1999; 99US-0151303P.  
 PR 31-AUG-1999; 99US-0151438P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152363P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 15-SEP-1999; 99US-0154018P.  
 PR 16-SEP-1999; 99US-0154039P.

PR 20-SEP-1999; 99US-0154779P.  
 PR 22-SEP-1999; 99US-0155139P.  
 PR 23-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155659P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159310P.  
 PR 14-OCT-1999; 99US-0159311P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 26-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161935P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 3.0%; Score 14; DB 3; Length 293;  
 Best Local Similarity 100.0%; Pred. No. 5.5e-06;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 264 PGQKVLVGGGIGG 277  
 DB 51 PGQKVLVGGGIGG 64

RESULT 10

AAG21652  
 ID AAG21652 standard; protein; 293 AA.

XX AC AAG21652;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 24284.

XX KW Protein identification; signal transduction pathway; metabolic pathway;  
 hybridisation assay; Genetic mapping; gene expression control; promoter;  
 termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-00301439.

XX PR 25-FEB-1999; 99US-0121825P.



PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132240P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134761P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135628P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136382P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140693P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.

PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145066P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 21-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145591P.  
PR 27-JUL-1999; 99US-0145591P.  
PR 27-JUL-1999; 99US-0145591P.  
PR 28-JUL-1999; 99US-0145591P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147392P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-SEP-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 26-SEP-1999; 99US-0156458P.  
PR 28-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.

PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158349P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159339P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 28-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 28-OCT-1999; 99US-0161992P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 3.0%; Score 14; DB 3; Length 293;

Best Local Similarity 100.0%; Pred. No. 5.5e-06; Mismatches 0; Indels 0; Gaps 0;

Matches 14; Conservative 0;

Qy 264 PGQKVLVDCGIGG 277  
 Db 51 PGQKVLVDCGIGG 64

RESULT 11  
 AAG45952  
 ID AAG45952 standard; protein; 293 AA.

AC AAG45952;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 57754.

XX Protein identification; signal transduction pathway; metabolic pathway;  
 XX hybridisation assay; genetic mapping; gene expression control; promoter;  
 XX termination sequence.

XX Arabidopsis thaliana.

XX EPI033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

XX 05-MAR-1999; 99US-0123180P.

XX 09-MAR-1999; 99US-0123548P.

XX 23-MAR-1999; 99US-0125788P.

XX 25-MAR-1999; 99US-0126264P.

XX 29-MAR-1999; 99US-0126785P.

XX 01-APR-1999; 99US-0127462P.

XX 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.  
 PR 16-APR-1999; 99US-0129845P.  
 PR 19-APR-1999; 99US-0130077P.  
 PR 21-APR-1999; 99US-0130449P.  
 PR 23-APR-1999; 99US-0130510P.  
 PR 23-APR-1999; 99US-0130891P.  
 PR 28-APR-1999; 99US-0131449P.  
 PR 30-APR-1999; 99US-0132048P.  
 PR 30-APR-1999; 99US-0132407P.  
 PR 30-APR-1999; 99US-0132484P.  
 PR 04-MAY-1999; 99US-0132485P.  
 PR 05-MAY-1999; 99US-0132486P.  
 PR 06-MAY-1999; 99US-0132487P.  
 PR 07-MAY-1999; 99US-0132863P.  
 PR 11-MAY-1999; 99US-0134256P.  
 PR 14-MAY-1999; 99US-0134218P.  
 PR 14-MAY-1999; 99US-0134219P.  
 PR 14-MAY-1999; 99US-0134221P.  
 PR 14-MAY-1999; 99US-0134370P.  
 PR 18-MAY-1999; 99US-0134768P.  
 PR 19-MAY-1999; 99US-0134941P.  
 PR 20-MAY-1999; 99US-0135124P.  
 PR 21-MAY-1999; 99US-0135353P.  
 PR 24-MAY-1999; 99US-0135629P.  
 PR 25-MAY-1999; 99US-0136021P.  
 PR 27-MAY-1999; 99US-0136392P.  
 PR 28-MAY-1999; 99US-0136782P.  
 PR 01-JUN-1999; 99US-0137222P.  
 PR 03-JUN-1999; 99US-0137528P.  
 PR 04-JUN-1999; 99US-0137502P.  
 PR 07-JUN-1999; 99US-0137724P.  
 PR 08-JUN-1999; 99US-0138094P.  
 PR 10-JUN-1999; 99US-0138540P.  
 PR 10-JUN-1999; 99US-0138847P.  
 PR 14-JUN-1999; 99US-0139119P.  
 PR 16-JUN-1999; 99US-0139452P.  
 PR 16-JUN-1999; 99US-0139453P.  
 PR 17-JUN-1999; 99US-0139492P.  
 PR 18-JUN-1999; 99US-0139454P.  
 PR 18-JUN-1999; 99US-0139455P.  
 PR 18-JUN-1999; 99US-0139456P.  
 PR 18-JUN-1999; 99US-0139457P.  
 PR 18-JUN-1999; 99US-0139458P.  
 PR 18-JUN-1999; 99US-0139459P.  
 PR 18-JUN-1999; 99US-0139460P.  
 PR 18-JUN-1999; 99US-0139461P.  
 PR 18-JUN-1999; 99US-0139462P.  
 PR 18-JUN-1999; 99US-0139463P.  
 PR 18-JUN-1999; 99US-0139750P.  
 PR 18-JUN-1999; 99US-0139763P.  
 PR 21-JUN-1999; 99US-0139817P.  
 PR 22-JUN-1999; 99US-0139899P.  
 PR 23-JUN-1999; 99US-0140353P.  
 PR 23-JUN-1999; 99US-0140354P.  
 PR 24-JUN-1999; 99US-0140695P.  
 PR 28-JUN-1999; 99US-0140823P.  
 PR 29-JUN-1999; 99US-0140991P.  
 PR 30-JUN-1999; 99US-0141287P.  
 PR 01-JUL-1999; 99US-0141842P.  
 PR 01-JUL-1999; 99US-0142154P.  
 PR 02-JUL-1999; 99US-0142055P.  
 PR 06-JUL-1999; 99US-0142390P.  
 PR 08-JUL-1999; 99US-0142803P.  
 PR 09-JUL-1999; 99US-0142920P.  
 PR 12-JUL-1999; 99US-0142977P.  
 PR 13-JUL-1999; 99US-0143542P.  
 PR 14-JUL-1999; 99US-0143624P.  
 PR 15-JUL-1999; 99US-0144005P.  
 PR 16-JUL-1999; 99US-0144085P.  
 PR 18-JUL-1999; 99US-0144086P.  
 PR 19-JUL-1999; 99US-0144325P.  
 PR 19-JUL-1999; 99US-0144331P.  
 PR 19-JUL-1999; 99US-0144332P.



PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132853P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135639P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136352P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139452P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 21-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140655P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140981P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144684P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.

```
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 22-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 25-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 3.0%; Score 14; DB 3; Length 296;
Best Local Similarity 100.0%; Pred. No. 5.5e-06; Mismatches 0; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 264 PGQKVLVDGCGIGG 277
DB 51 PGQKVLVDGCGIGG 64

RESULT 13
ABB93454
ID ABB93454 standard; protein; 314 AA.
AC ABB93454;
XX
XX 31-MAY-2002 (first entry)
XX Herbicidally active polypeptide SEQ ID NO 2665.
XX Herbicidal; plant; agriculture; herbicide.
XX Arabidopsis thaliana.
XX WO200210210-A2.
XX
XX PD 07-FEB-2002.
XX
XX PF 28-AUG-2001; 2001WO-EP009892.
XX
XX PF 28-AUG-2001; 2001WO-EP009892.
XX (FARB ) BAYER AG.
XX Tietjen K, Weidler M;
XX WPI; 2002-369010/31.
XX
XX PT Identifying plant target proteins for herbicidally active compounds,
XX comprising aligning and comparing nucleic acid or amino acid sequences
XX from plant with nucleic acid or amino acid sequences from non-plant
XX organisms.
XX
XX PS Claim 5; SEQ ID NO 2665; 261pp + Sequence Listing; English.
XX
XX CC The invention relates to identifying target proteins (ABB90790-ABB94016)
XX for herbicidally active compounds, comprising aligning and comparing
XX nucleic acid or amino acid sequences from plant with nucleic acid or
XX amino acid sequences from non-plant organisms using suitable search
XX parameters, where plant sequences having an E-value greater by a factor
XX of 3 than the E-value of most similar non-plant sequences are selected.
XX CC The polypeptides or nucleic acids encoding them are useful for
XX identifying modulators. The identified modulators are useful as
XX herbicides
```

```
XX SQ Sequence 314 AA;
Query Match 3.0%; Score 14; DB 5; Length 314;
Best Local Similarity 100.0%; Pred. No. 5.8e-06; Mismatches 0; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 264 PGQKVLVDGCGIGG 277
DB 72 PGQKVLVDGCGIGG 85

RESULT 14
AAG21651
ID AAG21651 standard; protein; 336 AA.
XX
XX AC AAG21651;
XX
XX DT 17-OCT-2000 (first entry)
XX
XX DB Arabidopsis thaliana protein fragment SEQ ID NO: 24283.
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridization assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX PN EP1033405-A2.
XX
XX PD 06-SEP-2000.
XX
XX PF 25-FEB-2000; 2000EP-00301439.
XX
XX PF 25-FEB-1999; 99US-0121825P.
XX PR 09-MAR-1999; 99US-0123180P.
XX PR 23-MAR-1999; 99US-0123548P.
XX PR 25-MAR-1999; 99US-0125788P.
XX PR 29-MAR-1999; 99US-0126264P.
XX PR 01-APR-1999; 99US-0126785P.
XX PR 08-APR-1999; 99US-0127462P.
XX PR 16-APR-1999; 99US-0128234P.
XX PR 19-APR-1999; 99US-0128714P.
XX PR 21-APR-1999; 99US-0129845P.
XX PR 23-APR-1999; 99US-0130077P.
XX PR 28-APR-1999; 99US-0130449P.
XX PR 30-APR-1999; 99US-0130510P.
XX PR 28-APR-1999; 99US-0130891P.
XX PR 30-APR-1999; 99US-0131449P.
XX PR 04-MAY-1999; 99US-0132048P.
XX PR 05-MAY-1999; 99US-0132407P.
XX PR 06-MAY-1999; 99US-0132484P.
XX PR 06-MAY-1999; 99US-0132485P.
XX PR 07-MAY-1999; 99US-0132486P.
XX PR 11-MAY-1999; 99US-0132487P.
XX PR 14-MAY-1999; 99US-0132863P.
XX PR 14-MAY-1999; 99US-0134256P.
XX PR 14-MAY-1999; 99US-0134218P.
XX PR 14-MAY-1999; 99US-0134219P.
XX PR 14-MAY-1999; 99US-0134221P.
XX PR 18-MAY-1999; 99US-0134370P.
XX PR 19-MAY-1999; 99US-0134768P.
XX PR 20-MAY-1999; 99US-0134941P.
XX PR 21-MAY-1999; 99US-0135124P.
XX PR 24-MAY-1999; 99US-0135333P.
XX PR 25-MAY-1999; 99US-0135629P.
XX PR 27-MAY-1999; 99US-0136021P.
XX PR 28-MAY-1999; 99US-0136392P.
XX PR 01-JUN-1999; 99US-0136782P.
XX PR 03-JUN-1999; 99US-0137222P.
XX PR 04-JUN-1999; 99US-0137528P.
XX PR 07-JUN-1999; 99US-0137502P.
XX PR 07-JUN-1999; 99US-0137724P.
```

PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 24-JUN-1999; 99US-0140823P.  
PR 26-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143543P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144684P.  
PR 21-JUL-1999; 99US-0144811P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 04-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.

PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149436P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149828P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149910P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161933P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match .....3.0% Score:14;...DB 3; Length 336;  
Best Local Similarity 100.0%; Pred. NO. 6.2e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 264 PCQKVLGVGGIGG 277  
DB 94 PCQKVLGVGGIGG 107

RESULT 15  
AAC45951  
ID AAC45951 standard; protein; 336 AA.  
AC AAC45951;  
XX  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 57753.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 05-MAY-1999; 99US-0132484P.  
PR 06-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 14-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137503P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 14-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 18-JUN-1999; 99US-0139817P.  
PR 21-JUN-1999; 99US-0139899P.  
PR 22-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 23-JUN-1999; 99US-0140695P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 02-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145132P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 04-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.



PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155559P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161933P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 3.0%; Score 14; DB 3; Length 336;  
Best Local Similarity 100.0%; Pred. No. 6.2e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 264 PGOKVLDVCGIGG 277  
Db 94 PGOKVLDVCGIGG 107

Search completed: July 26, 2004, 13:22:12  
Job time : 55 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2004, 13:21:12 / Search time 19 Seconds  
(without alignments)  
1285.215 Million cell updates/sec

Title: US-10-031-331B-40  
Perfect score: 473  
Sequence: 1 HTVDLTIEMMLDSQASLDL.....KLRRYAKGBQRMGLFVAKKK 473

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 389414 seqs, 51625971 residues

Word size : 0

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA:  
1: /cgm2\_6/ptodata/2/iaa/5A COMB.pap:  
2: /cgm2\_6/ptodata/2/iaa/5B COMB.pap:  
3: /cgm2\_6/ptodata/2/iaa/6A COMB.pap:  
4: /cgm2\_6/ptodata/2/iaa/6B COMB.pap:  
5: /cgm2\_6/ptodata/2/iaa/PCTUS COMB.pap:  
6: /cgm2\_6/ptodata/2/iaa/backfiles1.pap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	11	2.3	363	3	US-09-041-718-5
2	10	2.1	318	4	US-09-382-906A-2
3	8	1.7	151	4	US-09-690-454-150
4	8	1.7	243	3	US-09-181-958-1
5	8	1.7	244	4	US-09-328-352-5622
6	8	1.7	248	4	US-09-543-681A-4328
7	8	1.7	266	4	US-09-489-039A-8384
8	8	1.7	348	4	US-09-118-637A-4
9	8	1.7	352	4	US-09-489-039A-9039
10	8	1.7	376	3	US-09-041-718-2
11	8	1.7	501	1	US-08-331-394-4
12	8	1.7	501	1	US-08-250-858-4
13	8	1.7	501	1	US-08-446-915-4
14	8	1.7	501	2	US-08-744-139-4
15	8	1.7	501	4	US-08-779-599-4
16	8	1.7	501	5	PCT-US95-06639-4
17	7	1.5	36	4	US-08-216-592A-18
18	7	1.5	179	4	US-09-134-001C-4112
19	7	1.5	220	4	US-09-823-823-22
20	7	1.5	245	4	US-09-205-258-369
21	7	1.5	273	4	US-09-252-991A-28278
22	7	1.5	298	4	US-09-644-907B-8
23	7	1.5	336	4	US-09-543-681A-4420
24	7	1.5	348	4	US-09-252-991A-29436
25	7	1.5	355	4	US-09-644-907B-4
26	7	1.5	379	4	US-09-134-000C-4846
27	7	1.5	383	3	US-09-041-718-3

ALIGNMENTS

RESULT 1  
US-09-041-718-5  
; Sequence 5, Application US/09041718A  
; Patent No. 6225075  
; GENERAL INFORMATION:  
; APPLICANT: Bard, Martin  
; TITLE OF INVENTION: DNA encoding sterol methyltransferase  
; FILE REFERENCE: 740.003US1  
; CURRENT APPLICATION NUMBER: US/09/041,718A  
; CURRENT FILING DATE: 1998-03-13  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 363  
; TYPE: PRT  
; ORGANISM: Triticum atrivum  
US-09-041-718-5

Query Match 2.3%; Score 11; DB 3; Length 363;  
Best Local Similarity 100.0%; Pred. No. 0.012; 0; Indels 0; Gaps 0;  
Matches 11; Conservative 0; Mismatches 0;

QY 267 KVLVGGCGIGG 277  
DB 122 KVLVGGCGIGG 132

RESULT 2  
US-09-382-906A-2  
; Sequence 2, Application US/09382906A  
; Patent No. 6448475  
; GENERAL INFORMATION:  
; APPLICANT: Dellapenna, Dean  
; APPLICANT: Shintani, David  
; TITLE OF INVENTION: Manipulation of Tocopherol Levels in Transgenic Plants  
; FILE REFERENCE: 920905.90032  
; CURRENT APPLICATION NUMBER: US/09/382,906A  
; CURRENT FILING DATE: 1999-08-25  
; PRIOR APPLICATION NUMBER: 60/057,863  
; PRIOR FILING DATE: 1998-08-25  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 318  
; TYPE: PRT  
; ORGANISM: Synechocystis PCC6803  
US-09-382-906A-2

Query Match 2.1%; Score 10; DB 4; Length 318;  
Best Local Similarity 100.0%; Pred. No. 0.11;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 268 VLDVGGIGG 277  
| | | | |  
Db 97 VLDVGGIGG 106

## RESULT 3

US-09-690-454-150  
; Sequence 150, Application US/09690454

; Patent No. 6531447

; GENERAL INFORMATION:

; APPLICANT: Steven M. Ruben, et al.

; TITLE OF INVENTION: 32 Human Secreted Proteins

; FILE REFERENCE: P2006PI

; CURRENT APPLICATION NUMBER: US/09/690,454

; CURRENT FILING DATE: 2000-10-18

; PRIOR APPLICATION NUMBER: 09/189,144

; PRIOR FILING DATE: 1998-11-10

; PRIOR APPLICATION NUMBER: 60/044,039

; PRIOR FILING DATE: May 30, 1997

; PRIOR APPLICATION NUMBER: 60/048,093

; PRIOR FILING DATE: May 30, 1997

; PRIOR APPLICATION NUMBER: 60/044,190

; PRIOR FILING DATE: May 30, 1997

; PRIOR APPLICATION NUMBER: 60/050,935

; PRIOR FILING DATE: May 30, 1997

; PRIOR APPLICATION NUMBER: 60/048,101

; PRIOR FILING DATE: May 30, 1997

; PRIOR APPLICATION NUMBER: 60/048,356

; PRIOR FILING DATE: May 30, 1997

; PRIOR APPLICATION NUMBER: 60/056,250

; PRIOR FILING DATE: August 29, 1997

; PRIOR APPLICATION NUMBER: 60/056,296

; PRIOR FILING DATE: August 29, 1997

; PRIOR APPLICATION NUMBER: 60/056,293

; PRIOR FILING DATE: August 29, 1997

; NUMBER OF SEQ ID NOS: 229

; SOFTWARE: Patent In Ver. 2.0

; SEQ ID NO 150

; LENGTH: 151

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURES:

; NAME/KEY: SITE

; LOCATION: (123)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-690-454-150

Query Match 1.7%; Score 8; DB 4; Length 151;

Best Local Similarity 100.0%; Pred. No. 5.6;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 92 SPTLSFPP 99

| | | | |

Db 47 SPTLSFPP 54

| | | | |

RESULT 4

US-09-181-958-1

; Sequence 1, Application US/09181958

; Patent No. 6143507

; GENERAL INFORMATION:

; APPLICANT: Kehry, Marilyn R

; APPLICANT: Pullen, Steven S

; APPLICANT: Crute, James J

; TITLE OF INVENTION: High Throughput Compatible Assays for Receptor-TRAF

; TITLE OF INVENTION: Interactions

; FILE REFERENCE: 9 142 Nucl. Seq

; CURRENT APPLICATION NUMBER: US/09/181,958

; CURRENT FILING DATE: 1998-10-29

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: Patent In Ver. 2.0

US-09-181-958-1

Query Match 1.7%; Score 8; DB 4; Length 151;

Best Local Similarity 100.0%; Pred. No. 5.6;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 92 SPTLSFPP 99

| | | | |

Db 47 SPTLSFPP 54

| | | | |

; SEQ ID NO 1  
; LENGTH: 243  
; TYPE: PRT  
; ORGANISM: human  
US-09-181-958-1

Query Match 1.7%; Score 8; DB 3; Length 243;

Best Local Similarity 100.0%; Pred. No. 8.6;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 304 LRSIGLX 311

| | | | |

Db 53 LRSIGLX 60

| | | | |

RESULT 5

US-09-328-352-5622

; Sequence 5622, Application US/09328352

; Patent No. 6562958

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

; FILE REFERENCE: GTC99-03PA

; CURRENT APPLICATION NUMBER: US/09/328,352

; CURRENT FILING DATE: 1999-06-04

; NUMBER OF SEQ ID NOS: 8252

; SEQ ID NO 5622

; LENGTH: 244

; TYPE: PRT

; ORGANISM: Acinetobacter baumannii

US-09-328-352-5622

Query Match 1.7%; Score 8; DB 4; Length 244;

Best Local Similarity 100.0%; Pred. No. 8.6;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 KVLVGGCG 274

| | | | |

Db 60 KVLVGGCG 67

| | | | |

RESULT 6

US-09-543-681A-4328

; Sequence 4328, Application US/09543681A

; Patent No. 6605709

; GENERAL INFORMATION:

; APPLICANT: GARY BRETON

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS

; FILE REFERENCE: 2709.1002-001

; CURRENT APPLICATION NUMBER: US/09/543,681A

; CURRENT FILING DATE: 2000-04-05

; PRIOR APPLICATION NUMBER: US 60/128,706

; PRIOR FILING DATE: 1999-04-09

; NUMBER OF SEQ ID NOS: 8344

; SEQ ID NO 4328

; LENGTH: 248

; TYPE: PRT

; ORGANISM: Proteus mirabilis

US-09-543-681A-4328

Query Match 1.7%; Score 8; DB 4; Length 248;

Best Local Similarity 100.0%; Pred. No. 8.7;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 KVLVGGCG 274

| | | | |

Db 62 KVLVGGCG 69

| | | | |

RESULT 7

US-09-489-039A-8384

```
; Sequence 8384, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8384
; LENGTH: 266
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8384
Query Match 1.7%; Score 8; DB 4; Length 266;
Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 KVLVDVGGC 274
DB 83 KVLVDVGGC 90

RESULT 8
US-09-118-637A-4
; Sequence 4, Application US/09118637A
; Patent No. 6642434
; GENERAL INFORMATION:
; APPLICANT: DellaPenna, Dean
; APPLICANT: Shintani, David K.
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH TOCOPHEROL
; TITLE OF INVENTION: METHYLTRANSFERASE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 920905.90024
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-3166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 348 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-118-637A-4
Query Match 1.7%; Score 8; DB 4; Length 348;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 270 DVGGGIGG 277

; Sequence 8384, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8384
; LENGTH: 266
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8384
Query Match 1.7%; Score 8; DB 4; Length 266;
Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 KVLVDVGGC 274
DB 83 KVLVDVGGC 90

RESULT 8
US-09-118-637A-4
; Sequence 4, Application US/09118637A
; Patent No. 6642434
; GENERAL INFORMATION:
; APPLICANT: DellaPenna, Dean
; APPLICANT: Shintani, David K.
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH TOCOPHEROL
; TITLE OF INVENTION: METHYLTRANSFERASE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 920905.90024
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-3166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 348 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-118-637A-4
Query Match 1.7%; Score 8; DB 4; Length 348;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 270 DVGGGIGG 277

; Sequence 8384, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9039
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9039
Query Match 1.7%; Score 8; DB 4; Length 352;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 KVLVDVGGC 274
DB 209 KVLVDVGGC 216

RESULT 10
US-09-041-718-2
; Sequence 2, Application US/09041718A
; Patent No. 6225075
; GENERAL INFORMATION:
; APPLICANT: Bard, Martin
; TITLE OF INVENTION: DNA encoding sterol methyltransferase
; FILE REFERENCE: 740.003US1
; CURRENT APPLICATION NUMBER: US/09/041,718A
; CURRENT FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-041-718-2
Query Match 1.7%; Score 8; DB 3; Length 376;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 KVLVDVGGC 274
DB 125 KVLVDVGGC 132

RESULT 11
US-08-331-394-4
; Sequence 4, Application US/08331394
; Patent No. 5670319
; GENERAL INFORMATION:
; APPLICANT: Goeddel, David V.
; APPLICANT: Rothe, Mike
; TITLE OF INVENTION: Tumor Necrosis Factor
; TITLE OF INVENTION: Receptor-Associated Factors
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
```

CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/331.394  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/250858  
FILING DATE: 27-MAY-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Dreger, Ginger R.  
REGISTRATION NUMBER: 33,055  
REFERENCE/DOCKET NUMBER: 897P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-3216  
TELEFAX: 415/952-9881  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 501 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-331-394-4

Query Match 1.7%; Score 8; DB 1; Length 501;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 304 LERSIGLK 311  
DB 324 LERSIGLK 331

RESULT 12  
US-08-250-858-4  
Sequence 4, Application US/08250858  
Patent No. 5708142  
GENERAL INFORMATION:  
APPLICANT: Goedel, David V.  
TITLE OF INVENTION: Tumor Necrosis Factor Receptor-Associated Factors  
NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/250.858  
FILING DATE: 27-MAY-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Dreger, Ginger R.  
REGISTRATION NUMBER: 33,055  
REFERENCE/DOCKET NUMBER: 897.1  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/225-3216  
TELEFAX: 415/952-9881  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 501 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-250-858-4

Query Match 1.7%; Score 8; DB 1; Length 501;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 304 LERSIGLK 311  
DB 324 LERSIGLK 331

RESULT 13  
US-08-446-915-4  
Sequence 4, Application US/08446915  
Patent No. 5741667  
GENERAL INFORMATION:  
APPLICANT: Goedel, David V.  
TITLE OF INVENTION: Tumor Necrosis Factor Receptor-Associated Factors  
NUMBER OF SEQUENCES: 66  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/446.915  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/250858  
FILING DATE: 27-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/331394  
FILING DATE: 28-OCT-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Dreger, Ginger R.  
REGISTRATION NUMBER: 33,055  
REFERENCE/DOCKET NUMBER: 897P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-3216  
TELEFAX: 415/952-9881  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 501 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-446-915-4

Query Match 1.7%; Score 8; DB 1; Length 501;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 304 LERSIGLK 311  
DB 324 LERSIGLK 331

RESULT 14

US-08-744-139-4  
; Sequence 4, Application US/08744139  
; Patent No. 5869612  
; GENERAL INFORMATION:  
; APPLICANT: Goedel, David V  
; APPLICANT: Rothe, Mike  
; TITLE OF INVENTION: TUMOR NECROSIS FACTOR RECEPTOR-ASSOCIATED FACTORS  
; NUMBER OF SEQUENCES: 59  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/744,139  
; FILING DATE: 31-Oct-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION NUMBER: 08/250858  
; FILING DATE: 05/27/1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dregler, Ginger R.  
; REGISTRATION NUMBER: 33,055  
; REFERENCE/DOCKET NUMBER: P0897C1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-3216  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 501 amino acids  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
US-08-744-139-4

Query Match 1.7%; Score 8; DB 2; Length 501;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 304 LERSIGLK 311  
DB 324 LERSIGLK 331

RESULT 15

US-08-779-599-4  
; Sequence 4, Application US/08779599  
; Patent No. 6500922  
; GENERAL INFORMATION:  
; APPLICANT: Goedel, David V.  
; APPLICANT: Rothe, Mike  
; TITLE OF INVENTION: Tumor Necrosis Factor Receptor-Associated Factors  
; NUMBER OF SEQUENCES: 59  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/779,599  
; FILING DATE: 07-Jan-1997  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dregler, Ginger R.  
; REGISTRATION NUMBER: 33,055  
; REFERENCE/DOCKET NUMBER: P0897C2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-3216  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 501 amino acids  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
US-08-779-599-4

Query Match 1.7%; Score 8; DB 4; Length 501;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 304 LERSIGLK 311  
DB 324 LERSIGLK 331

Search completed: July 26, 2004, 13:24:28  
Job time : 20 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2004, 13:23:33 ; Search time 45 Seconds  
(without alignments)  
3291.599 Million cell updates/sec

Title: US-10-031-331b-40

Perfect score: 473

Sequence: 1 HTVDLTTEAMLDSQASDLD.....KLRTAKGEQWGLFVAKK 473

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1288442 seqs, 313154207 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1288442

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	5.5	219	12	US-10-425-114-53753
2	26	5.5	399	12	US-10-425-114-72568
3	26	5.5	513	12	US-10-425-114-37470
4	25	5.3	314	12	US-10-425-114-69972
5	23	4.9	331	12	US-10-425-114-44987
6	21	4.4	168	14	US-10-424-599-759-50
7	21	4.4	491	14	US-10-424-599-759-82
8	18	3.8	183	12	US-10-424-599-240673
9	17	3.6	133	16	US-10-437-963-110034
10	17	3.6	501	12	US-10-424-599-172601
11	17	3.6	530	16	US-10-437-963-155711
12	16	3.4	196	12	US-10-424-599-197735
13	16	3.4	210	12	US-10-424-599-197735
14	16	3.4	231	12	US-10-424-599-197733
15	16	3.4	246	12	US-10-424-599-197731

16	16	3.4	271	16	US-10-437-963-110035	Sequence 110035,
17	16	3.4	324	12	US-10-424-599-246283	Sequence 246283,
18	16	3.4	344	12	US-10-425-114-36573	Sequence 36573, A
19	16	3.4	344	12	US-10-425-114-36660	Sequence 36660, A
20	16	3.4	377	12	US-10-425-114-60338	Sequence 60338, A
21	16	3.4	408	12	US-10-424-599-197736	Sequence 197736,
22	14	3.0	120	16	US-10-437-963-155679	Sequence 155679,
23	13	2.7	237	12	US-10-424-599-257495	Sequence 257495,
24	11	2.3	58	12	US-10-424-599-201471	Sequence 201471,
25	11	2.3	138	16	US-10-437-963-167513	Sequence 167513,
26	11	2.3	202	10	US-09-768-235B-22	Sequence 22, Appl
27	11	2.3	285	15	US-10-425-134A-14	Sequence 14, Appl
28	11	2.3	338	12	US-10-425-114-41948	Sequence 41948, A
29	11	2.3	344	9	US-09-779-144A-7	Sequence 7, Appl
30	11	2.3	345	12	US-10-425-114-39343	Sequence 39343, A
31	11	2.3	345	16	US-10-437-963-147769	Sequence 147769,
32	11	2.3	351	12	US-10-425-114-56209	Sequence 56209, A
33	11	2.3	351	12	US-10-425-114-65586	Sequence 65586, A
34	11	2.3	358	12	US-10-425-114-55779	Sequence 55779, A
35	11	2.3	358	12	US-10-425-114-59573	Sequence 59573, A
36	11	2.3	362	16	US-10-437-963-147768	Sequence 147768,
37	11	2.3	384	12	US-10-425-114-39121	Sequence 39121, A
38	11	2.3	383	16	US-10-437-963-184331	Sequence 184331,
39	11	2.3	409	16	US-10-437-963-167512	Sequence 167512,
40	11	2.3	412	12	US-10-425-114-42141	Sequence 42141, A
41	10	2.1	309	14	US-10-219-810-48	Sequence 48, Appl
42	10	2.1	318	14	US-10-219-810-46	Sequence 46, Appl
43	10	2.1	318	14	US-10-380-132-22	Sequence 22, Appl
44	10	2.1	318	16	US-10-471-243-20	Sequence 20, Appl
45	10	2.1	437	16	US-10-602-268-11	Sequence 11, Appl

## ALIGNMENTS

## RESULT 1

US-10-425-114-53753  
; Sequence 53753, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 53753  
; LENGTH: 219  
; TYPE: PRT  
; ORGANISM: Gossypium hirsutum  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3829-031-F4\_FLI.pep  
US-10-425-114-53753

Query Match 5.5% Score 26; DB 12; Length 219;  
Best Local Similarity 100.0%; Pred. No. 2.3e-17;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 326 YPDNSFDVIYSRDTILHIQDKPALFR 351  
DB 72 YPDNSFDVIYSRDTILHIQDKPALFR 97

## RESULT 2

US-10-425-114-72568  
; Sequence 72568, Application US/10425114  
; Publication No. US20040034888A1



GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; NUMBER OF SEQ ID NOS: 2003-04-28  
; CURRENT FILING DATE: 2003-04-28  
; SEQ ID NO 72568  
; LENGTH: 399  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB23-061-D6\_FLI.pep  
US-10-425-114-72568

Query Match 5.5%; Score 26; DB 12; Length 399;  
Best Local Similarity 100.0%; Pred. No. 3.9e-17;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 YPDSFVYISRDITLHIQDKPALFR 351  
DB 252 YPDSFVYISRDITLHIQDKPALFR 277

RESULT 3  
US-10-425-114-37470  
; Sequence 37470, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; NUMBER OF SEQ ID NOS: 2003-04-28  
; CURRENT FILING DATE: 2003-04-28  
; SEQ ID NO 37470  
; LENGTH: 513  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB22-074-F3\_FLI.pep  
US-10-425-114-37470

Query Match 5.5%; Score 26; DB 12; Length 513;  
Best Local Similarity 100.0%; Pred. No. 5e-17;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 YPDSFVYISRDITLHIQDKPALFR 351  
DB 366 YPDSFVYISRDITLHIQDKPALFR 391

RESULT 4  
US-10-425-114-69972  
; Sequence 69972, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E

APPLICANT: Tabaska, Jack E  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53313)B  
CURRENT APPLICATION NUMBER: US/10/425,114  
NUMBER OF SEQ ID NOS: 2003-04-28  
CURRENT FILING DATE: 2003-04-28  
SEQ ID NO 69972  
LENGTH: 314  
TYPE: PRT  
ORGANISM: Zea mays subsp. mexicana  
FEATURE:  
OTHER INFORMATION: Clone ID: UC-ZMROTEOSINTE108E03\_FLI.pep  
US-10-425-114-69972

Query Match 5.3%; Score 25; DB 12; Length 314;  
Best Local Similarity 100.0%; Pred. No. 3.2e-16;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 259 MLDLKPQKVLVDGCGIGGDFYMA 283  
DB 101 MLDLKPQKVLVDGCGIGGDFYMA 125

RESULT 5  
US-10-425-114-44987  
; Sequence 44987, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; NUMBER OF SEQ ID NOS: 2003-04-28  
; CURRENT FILING DATE: 2003-04-28  
; SEQ ID NO 44987  
; LENGTH: 331  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700211781\_FLI.pep  
US-10-425-114-44987

Query Match 4.9%; Score 23; DB 12; Length 331;  
Best Local Similarity 100.0%; Pred. No. 3.5e-14;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 262 LKPGQKVLVDGCGIGGDFYMAE 284  
DB 121 LKPGQKVLVDGCGIGGDFYMAE 143

RESULT 6  
US-10-149-759-50  
; Sequence 50, Application US/10149759  
; Publication No. US20030157592A1  
; GENERAL INFORMATION:  
; APPLICANT: Lerchl, Jens  
; APPLICANT: Renz, Andreas  
; APPLICANT: Ehrhardt, Thomas  
; APPLICANT: Reindl, Andreas  
; APPLICANT: Cirpus, Petra  
; TITLE OF INVENTION: Moss Genes from Physcomitrella patens encoding proteins  
; TITLE OF INVENTION: Involved in the synthesis of tocopherols and  
; FILE REFERENCE: BASF/NAE 1333/99 PCT/US

; CURRENT APPLICATION NUMBER: US/10/149,759  
; CURRENT FILING DATE: 2002-10-17  
; PRIOR APPLICATION NUMBER: PCT/EP/00/12698  
; PRIOR FILING DATE: 2000-12-14  
; NUMBER OF SEQ ID NOS: 82  
; SOFTWARE: WordPerfect version 6.1  
; SEQ ID NO 50  
; LENGTH: 168  
; TYPE: PRT  
; ORGANISM: Physcomitrella patens  
US-10-149-759-50

Query Match 4.4%; Score 21; DB 14; Length 168;  
Best Local Similarity 100.0%; Pred. No. 1.9e-12;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 330 SFDVIYSRDTLHIQDKPALF 350  
|||  
DB 24 SFDVIYSRDTLHIQDKPALF 44  
|||

## RESULT 7

US-10-149-759-82  
; Sequence 82, Application US/10149759  
; Publication No. US20030157592A1  
; GENERAL INFORMATION:  
; APPLICANT: Lerchl, Jens  
; APPLICANT: Renz, Andreas  
; APPLICANT: Ehrhardt, Thomas  
; APPLICANT: Reindl, Andreas  
; APPLICANT: Cirpug, Petra  
; TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins  
; TITLE OF INVENTION: Involved in the synthesis of tocopherols and  
; FILE REFERENCE: BASF/NAB 1333/99 PCT/US  
; CURRENT APPLICATION NUMBER: US/10/149,759  
; CURRENT FILING DATE: 2002-10-17  
; PRIOR APPLICATION NUMBER: PCT/EP/00/12698  
; PRIOR FILING DATE: 2000-12-14  
; NUMBER OF SEQ ID NOS: 82  
; SOFTWARE: WordPerfect version 6.1  
; SEQ ID NO 82  
; TYPE: PRT  
; ORGANISM: Physcomitrella patens  
US-10-149-759-82

Query Match 4.4%; Score 21; DB 14; Length 491;  
Best Local Similarity 100.0%; Pred. No. 5.1e-12;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 PGGYIFFRSCFHQSGDHKK 152  
|||  
DB 148 PGGYIFFRSCFHQSGDHKK 168  
|||

## RESULT 8

US-10-424-599-240673  
; Sequence 240673, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic, David K  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 240673  
; LENGTH: 183  
; TYPE: PRT

; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(183)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_59355C.1.pep  
US-10-424-599-240673

Query Match 3.8%; Score 18; DB 12; Length 183;  
Best Local Similarity 100.0%; Pred. No. 2.2e-09;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 187 SKCIGAYVRNKKQNOI 204  
|||  
DB 96 SKCIGAYVRNKKQNOI 113  
|||

## RESULT 9

US-10-437-963-110034  
; Sequence 110034, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 110034  
; LENGTH: 133  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_14136C.1.pep  
US-10-437-963-110034

Query Match 3.6%; Score 17; DB 16; Length 133;  
Best Local Similarity 100.0%; Pred. No. 1.6e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 378 EFAAYIKQGYDLHDVK 394  
|||  
DB 39 EFAAYIKQGYDLHDVK 55  
|||

## RESULT 10

US-10-424-599-172601  
; Sequence 172601, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic, David K  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 172601  
; LENGTH: 501  
; TYPE: PRT



; NAME/KEY: unsure  
; LOCATION: (1)..(231)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_20579C.1.pep  
US-10-424-599-197731

Query Match 3.4%; Score 16; DB 12; Length 231;  
Best Local Similarity 100.0%; Pred. No. 2.7e-07;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 262 LKPGQKVLVDVCGIGG 277  
||| ||||| ||||| |||||  
DB 141 LKPGQKVLVDVCGIGG 156

RESULT 15  
US-10-424-599-197731  
; Sequence 197731, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 197731  
; LENGTH: 246  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(246)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_20577C.1.pep  
US-10-424-599-197731

Query Match 3.4%; Score 16; DB 12; Length 246;  
Best Local Similarity 100.0%; Pred. No. 2.9e-07;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 262 LKPGQKVLVDVCGIGG 277  
||| ||||| ||||| |||||  
DB 138 LKPGQKVLVDVCGIGG 153

Search completed: July 26, 2004, 13:29:12  
Job time : 45 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2004, 13:20:17 ; Search time 17 Seconds  
(without alignments)  
2676.388 Million cell updates/sec

Title: US-10-031-331B-40  
Perfect score: 473  
Sequence: 1 HTVDLTIANMQLDSQSLD.....KURRTAKGQRGLFVAKKK 473

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283366 seqs, 96191526 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR-78.\*

1: PIR1.\*

2: PIR2.\*

3: PIR3.\*

4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	5.5	374	2	protein T1N15.23 (
2	26	5.5	555	2	hypothetical prote
3	16	3.4	367	2	probable sterol 24
4	12	2.5	346	2	sterol 24-C-methyl
5	11	2.3	344	2	sterol 24-C-methyl
6	11	2.3	344	2	sterol 24-C-methyl
7	11	2.3	363	2	probable sterol 24
8	10	2.1	318	2	hypothetical prote
9	9	1.9	280	2	gamma-cocopherol m
10	9	1.9	330	2	hypothetical prote
11	9	1.9	437	2	hypothetical prote
12	9	1.9	761	2	adenylate cyclase
13	9	1.9	1839	1	3-demethylubiquino
14	8	1.7	209	1	probable IS1560 tr
15	8	1.7	228	2	3-demethylubiquino
16	8	1.7	240	1	hypothetical prote
17	8	1.7	240	2	hypothetical prote
18	8	1.7	240	2	hypothetical prote
19	8	1.7	242	2	3-demethylubiquino
20	8	1.7	242	2	3-demethylubiquino
21	8	1.7	257	2	hypothetical prote
22	8	1.7	283	2	conserved hypotet
23	8	1.7	289	2	probable methyltra
24	8	1.7	342	2	conserved hypotet
25	8	1.7	343	2	ribosomal RNA smal
26	8	1.7	343	2	probable enzyme V
27	8	1.7	343	2	tRNA (guanine-N2-)
28	8	1.7	348	2	gamma-cocopherol m
29	8	1.7	379	2	probable DELTA(24)

30 8 1.7 416 2 A70393 hypothetical prote  
31 8 1.7 495 2 T27936 hypothetical prote  
32 8 1.7 501 2 T61512 TNF receptor assoc  
33 8 1.7 501 2 S56163 tumor necrosis fac  
34 8 1.7 872 2 S49541 cellulase - Cellul  
35 7 1.5 71 2 C69957 hypothetical prote  
36 7 1.5 78 2 F82823 hypothetical prote  
37 7 1.5 98 2 T15767 hypothetical prote  
38 7 1.5 100 2 H71132 hypothetical prote  
39 7 1.5 111 2 B45824 flal protein - Bac  
40 7 1.5 111 2 B25159 13K sin operon hyp  
41 7 1.5 128 1 FEHSX ferredoxin [2Fe-2S  
42 7 1.5 128 2 S19497 hypothetical prote  
43 7 1.5 129 1 FEHS ferredoxin [2Fe-2S  
44 7 1.5 129 2 T43917 ferredoxin [2Fe-2S  
45 7 1.5 129 2 B84380 ferredoxin import

## ALIGNMENTS

### RESULT 1

P96525  
protein T1N15.23 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001

C/Accession: P96525

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: P96525

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-374 <STO>

A/Cross-references: GB:A8005173; MID:g8778697; PIDN:AAF79705.1; GSPDB:GN00141

C/Genetic:

A:Gene: T1N15.23

A:Map position: 1

Query Match

Best Local Similarity 5.5%; Score 26; DB 2; Length 374;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 326 YPDNSFDVYSRDTLHTQDKPALFR 351

Db 198 YPDNSFDVYSRDTLHTQDKPALFR 223

### RESULT 2

H96762

hypothetical protein P6D5.1 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001

C/Accession: H96762

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: H96762  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-555 <STO>  
A:Cross-references: GB:AE005173; NID:g10092368; PIDN:AAG12776.1; GSPDB:GN00141  
C:Genetics:  
A:Gene: F6D5.1  
A:Map position: 1

Query Match 5.5%; Score 26; DB 2; Length 555;  
Best Local Similarity 100.0%; Pred. No. 4.4e-18;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 259 MLDLKPQKVLVDVCGIGSGDFYMAE 284  
DB 341 MLDLKPQKVLVDVCGIGSGDFYMAE 366

RESULT 3  
T06780  
sterol 24-C-methyltransferase (SC 2.1.1.41) - soybean  
N:Alternate names: S-adenosyl-L-methionine:delta24-sterol-C-methyltransferase  
C:Species: Glycine max (soybean)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 03-Jun-2002  
C:Accession: T06780  
R:Shi, J.; Gonzales, R.A.; Bhattacharyya, M.K.  
J. Biol. Chem. 271, 9384-9389, 1996  
A:Title: Identification and characterization of an S-adenosyl-L-methionine: delta 24-sterol  
A:Reference number: Z15807; MUID:96199190; PMID:8621604  
A:Accession: T06780  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-367 <SHI>  
A:Cross-references: EMBL:U43683; NID:g1399379; PIDN:AAB04057.1; PID:g1399380  
A:Experimental source: Cultivar Williams 82; etiolated hypocotyls  
C:Superfamily: 24-sterol C-methyltransferase; bioC homology  
C:Keywords: methyltransferase; S-adenosylmethionine  
F:122-226/Domain: bioC homology <BIOC>

Query Match 3.4%; Score 16; DB 2; Length 367;  
Best Local Similarity 100.0%; Pred. No. 5.8e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 262 LKPGQKVLVDVCGIGG 277  
DB 119 LKPGQKVLVDVCGIGG 134

RESULT 4  
T10173  
sterol 24-C-methyltransferase (SC 2.1.1.41) - castor bean  
N:Alternate names: S-adenosyl-methionine-sterol-C-methyltransferase  
C:Species: Ricinus communis (castor bean)  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 03-Jun-2002  
C:Accession: T10173  
R:Bouvier-Nave, P.; Hesselstein, T.; Desprez, T.; Benveniste, P.  
Eur. J. Biochem. 246, 518-529, 1997  
A:Title: Identification of cDNAs encoding sterol methyl-transferases involved in the sec  
A:Reference number: Z15116; MUID:97352551; PMID:9208946  
A:Accession: T10173  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-346 <BOU>  
A:Cross-references: EMBL:U81313; NID:g2246457; PIDN:AAB62812.1; PID:g2246458  
A:Experimental source: strain Baker 296  
C:Superfamily: 24-sterol C-methyltransferase; bioC homology  
C:Keywords: methyltransferase; S-adenosylmethionine

Query Match 2.5%; Score 12; DB 2; Length 346;  
Best Local Similarity 100.0%; Pred. No. 0.00072;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 266 QKVLVDVCGIGG 277

DB 102 QKVLVDVCGIGG 113

RESULT 5  
T01572  
sterol 24-C-methyltransferase (EC 2.1.1.41) - maize  
N:Alternate names: (S)-adenosyl-L-methionine:delta 24-sterol methyltransferase  
C:Species: Zea mays (maize)  
C:Date: 19-Feb-1999 #sequence\_revision 19-Feb-1999 #text\_change 03-Jun-2002  
C:Accession: T01572  
R:Tong, Y.; Nee, W.D.  
submitted to the EMBL Data Library, January 1998  
A:Reference number: Z14350  
A:Accession: T01572  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-344 <TON>  
A:Cross-references: EMBL:AF045570; NID:g2909845; PIDN:AAC04265.1; PID:g2909846  
A:Experimental source: strain B73  
C:Genetics:  
A:Gene: SMT  
C:Superfamily: 24-sterol C-methyltransferase; bioC homology  
C:Keywords: methyltransferase; S-adenosylmethionine  
F:101-205/Domain: bioC homology <BIOC>

Query Match 2.3%; Score 11; DB 2; Length 344;  
Best Local Similarity 100.0%; Pred. No. 0.0076;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 KVLVDVCGIGG 277  
DB 103 KVLVDVCGIGG 113

RESULT 6  
T04138  
sterol 24-C-methyltransferase (EC 2.1.1.41) ESMT1, endosperm - maize  
N:Alternate names: C-24 sterol methyltransferase  
C:Species: Zea mays (maize)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 03-Jun-2002  
C:Accession: T04138  
R:Grebenok, R.J.; Galbraith, D.W.; Dellepenna, D.  
Plant Mol. Biol. 34, 891-896, 1997  
A:Title: Characterization of zeamays endosperm C-24 sterol methyltransferase - one of  
A:Reference number: Z09668; MUID:97435974; PMID:9290641  
A:Accession: T04138  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-344 <GRS>  
A:Cross-references: EMBL:U79669; NID:g1899059; PIDN:AAB70886.1; PID:g1899060  
A:Experimental source: endosperm  
C:Genetics:  
A:Gene: ESMT1  
C:Function:  
A:Description: methyltransferase  
C:Superfamily: 24-sterol C-methyltransferase; bioC homology  
C:Keywords: methyltransferase; S-adenosylmethionine  
F:101-205/Domain: bioC homology <BIOC>

Query Match 2.3%; Score 11; DB 2; Length 344;  
Best Local Similarity 100.0%; Pred. No. 0.0076;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 KVLVDVCGIGG 277  
DB 103 KVLVDVCGIGG 113

RESULT 7  
T06795  
probable sterol 24-C-methyltransferase (EC 2.1.1.41) - wheat  
N:Alternate names: delta-24-sterol methyltransferase

C:Species: Triticum aestivum (common wheat)  
 C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 03-Jun-2002  
 C:Accession: T06795  
 R:Subramaniam, K.; Ueng, P.P.  
 A:Reference number: 215820  
 A:Reference number: 215820  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-363 <SUB>  
 A:Cross-references: EMBL:U60754; NID:gl706964; PIDN:AB37769.1; PID:gl706965  
 C:Genetics:

Query Match 2.3%; Score 11; DB 2; Length 363;  
 Best Local Similarity 100.0%; Pred. No. 0.008;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 KVLVGGGIGG 277  
 |||||  
 DB 122 KVLVGGGIGG 132

## RESULT 8

S76226  
 hypothetical protein - Synchocystis sp. (strain PCC 6803)  
 C:Species: Synchocystis sp.  
 A:Variety: PCC 6803  
 C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999  
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda  
 DNA Res. 3, 109-136, 1996  
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synchocystis  
 A:Reference number: S74322; MUID:97061201; PMID:8905231  
 A:Accession: S76226  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-318 <KAN>  
 A:Cross-references: EMBL:D90914; GB:AB001339; NID:gl653477; PIDN:BAAL8485.1; PID:d101924  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
 C:Superfamily: bioC homology  
 F:94-195/Domain: bioC homology <BIOC>

Query Match 2.1%; Score 10; DB 2; Length 318;  
 Best Local Similarity 100.0%; Pred. No. 0.076;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 268 VLDVGGGIGG 277  
 |||||  
 DB 97 VLDVGGGIGG 106

## RESULT 9

AE2031  
 gamma-tocopherol methyltransferase [imported] - Nostoc sp. (strain PCC 7120)  
 C:Species: Nostoc sp. PCC 7120  
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
 C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
 C:Accession: AE2031  
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi  
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S  
 DNA Res. 8, 205-213, 2001  
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
 A:Reference number: AB1807; MUID:21595285; PMID:11759840  
 A:Accession: AE2031  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-280 <KUR>

A:Cross-references: GB:BA000019; PIDN:BA073502.1; PID:gl7130893; GSPDB:GN00179  
 A:Experimental source: strain PCC 7120  
 C:Genetics:  
 A:Gene: alr1803

Query Match 1.9%; Score 9; DB 2; Length 280;  
 Best Local Similarity 100.0%; Pred. No. 0.72;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 269 LDVGGGIGG 277  
 |||||  
 DB 67 LDVGGGIGG 75

## RESULT 10

AC2071  
 hypothetical protein all2121 [imported] - Nostoc sp. (strain PCC 7120)  
 C:Species: Nostoc sp. PCC 7120  
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
 C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
 C:Accession: AC2071  
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi  
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S  
 DNA Res. 8, 205-213, 2001  
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
 A:Reference number: AB1807; MUID:21595285; PMID:11759840  
 A:Accession: AC2071  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-330 <KUR>  
 A:Cross-references: GB:BA000019; PIDN:BA073820.1; PID:gl7131212; GSPDB:GN00179  
 A:Experimental source: strain PCC 7120  
 C:Genetics:  
 A:Gene: all2121

Query Match 1.9%; Score 9; DB 2; Length 330;  
 Best Local Similarity 100.0%; Pred. No. 0.83;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 269 LDVGGGIGG 277  
 |||||  
 DB 97 LDVGGGIGG 105

## RESULT 11

T29330  
 hypothetical protein F54D11.1 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T29330  
 R:Pauley, A.; Gattung, S.  
 submitted to the EMBL Data Library, July 1996  
 A:Description: The sequence of C. elegans cosmid F54D11.  
 A:Reference number: Z20606  
 A:Accession: T29330  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-437 <PAU>  
 A:Cross-references: EMBL:U64834; PIDN:AB04824.1; GSPDB:GN00023; CESP:F54D11.1  
 A:Experimental source: strain Bristol N2; clone F54D11  
 C:Genetics:  
 A:Gene: CESP:F54D11.1  
 A:Map position: 5  
 A:Introns: 36/2; 76/3; 108/1; 315/3; 398/3

Query Match 1.9%; Score 9; DB 2; Length 437;  
 Best Local Similarity 100.0%; Pred. No. 1.1;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 357 LKPGGKVI 365  
 |||||  
 DB 318 LKPGGKVI 326



## RESULT 12

H65083  
glycolate oxidase subunits GlcE and GlcF - Escherichia coli (strain K-12)  
C:Species: Escherichia coli  
C>Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 01-Mar-2002  
C:Accession: H65083  
R:Blattner, P.R.; Maun, B.; Shao, Y.  
A: Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617; PMID:9278503  
A:Accession: H65083  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA

A:Residues: 1-761 <BLAT>  
A:Cross-references: GB:A0000380; GB:U00096; NID:91789344; PIDN:AAC76014.1; PID:91789350;  
A:Experimental source: strain K-12, substrain MG1655  
C:Superfamily: unassigned ferredoxin 2[4Fe-4S]-related proteins; ferredoxin 2[4Fe-4S] hcd  
P:372-447/Domain: ferredoxin 2[4Fe-4S] homology <FER2>

Query Match 1.9%; Score 9; DB 2; Length 761;  
Best Local Similarity 100.0%; Pred. No. 1.8;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 VNEYGQMLK 401  
|||||  
DB 609 VNEYGQMLK 617

## RESULT 13

OYBVK  
adenylate cyclase (EC 4.6.1.1) - yeast (Saccharomyces kluyveri)  
N:Alternate names: adenylyl cyclase  
C:Species: Saccharomyces kluyveri  
C>Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 21-Jan-2000  
C:Accession: JQ1145; S14464  
R:Young, D.; O'Neill, K.; Broek, D.; Wigler, M.  
Gene 102, 129-132, 1991  
A:Title: The adenylyl cyclase-encoding gene from Saccharomyces kluyveri.  
A:Reference number: JQ1145; MUID:91323718; PMID:1864503  
A:Accession: JQ1145  
A:Molecule type: DNA

A:Residues: 1-1839 <YOU>  
A:Cross-references: EMBL:X56042; NID:94856; PIDN:CAA39513.1; PID:94857  
C:Genetics:

A:Gene: CYR1  
C:Superfamily: yeast adenylate cyclase; leucine-rich alpha-2-glycoprotein repeat homolog  
C:Keywords: CAMP biosynthesis; duplication; phosphorus-oxygen lyase  
P:1006-1029/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>  
P:1479-1564/Domain: yeast adenylate cyclase catalytic domain homology <YACC>

Query Match 1.9%; Score 9; DB 1; Length 1839;  
Best Local Similarity 100.0%; Pred. No. 3.8;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 LDFIESAIK 71  
|||||  
DB 650 LDFIESAIK 658

## RESULT 14

S32628  
3-demethylubiquinone-9 3-O-methyltransferase (EC 2.1.1.64) - Salmonella typhimurium (fra  
N:Alternate names: 2-octaprenyl-3-methyl-5-hydroxy-6-methoxy-1,4-benzoquinone methyltran  
C:Species: Salmonella typhimurium  
C>Date: 28-May-1999 #sequence\_revision 28-May-1999 #text\_change 05-May-2000  
C:Accession: S32628  
R:Jordan, A.

submitted to the EMBL Data Library, March 1993

A:Reference number: S32628  
A:Accession: S32628  
A:Molecule type: DNA

A:Residues: 1-209 <JOR>

A:Cross-references: EMBL:X72948; NID:9510348; PIDN:CAA51451.1; PID:9295895  
C:Superfamily: 3-demethylubiquinone-9 3-O-methyltransferase; bioC homology  
C:Keywords: methyltransferase; S-adenosylmethionine; ubiquinone biosynthesis  
P:24-126/Domain: bioC homology <BIOC>

Query Match 1.7%; Score 8; DB 1; Length 209;  
Best Local Similarity 100.0%; Pred. No. 5.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 KVLVDVGGC 274  
|||||  
DB 26 KVLVDVGGC 33

## RESULT 15

P70624  
probable IS1560 truncated transposase Rv1035c - Mycobacterium tuberculosis (strain H37RV  
C:Species: Mycobacterium tuberculosis  
C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 02-Sep-2000  
C:Accession: F70624

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.  
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sultston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: P70624

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-228 <COL>

A:Cross-references: GB:Z92539; GB:AL123456; NID:93261714; PIDN:CAB06844.1; PID:e304621;

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: Rv1035c

Query Match 1.7%; Score 8; DB 2; Length 228;  
Best Local Similarity 100.0%; Pred. No. 6.4;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 RPTGELAE 55  
|||||  
DB 121 RPTGELAE 128

Search completed: July 26, 2004, 13:23:58

Job time: 18 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2004, 13:16:27 ; Search time 13 Seconds  
(without alignments)  
1894.553 Million cell updates/sec

Title: US-10-031-331B-40

Perfect score: 473

Sequence: 1 HTVDLTTEAMLDQASOLD.....KLRTAKGEQWGLFVAKKK 473

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 141681 seqs, 52070155 residues

Word size : 0

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	9.5	494	PEAM_SPIOL	Q9m571 spinacia ol
2	26	5.5	475	FEM2_ARATH	Q944h0 arabidopsis
3	26	5.5	490	FEM3_ARATH	Q946b9 arabidopsis
4	26	5.5	491	FEM1_ARATH	Q94744 arabidopsis
5	9	1.9	365	1D12_METKA	Q8x999 methanopyru
6	9	1.9	407	GLCF_ECOLI	P52074 escherichia
7	9	1.9	1839	CYAA_SAKL	P23466 saccharomyc
8	8	1.7	232	UBIG_CHRVO	Q7nz91 chromobacte
9	8	1.7	232	UBIG_PSEPK	Q88m10 pseudomonas
10	8	1.7	232	UBIG_PSEPM	Q885t9 pseudomonas
11	8	1.7	235	UBIG_VIRPA	Q87nd5 vibrio para
12	8	1.7	235	UBIG_VIRVU	Q8d8e0 vibrio vuln
13	8	1.7	235	UBIG_VIRVU	Q7mm27 vibrio vuln
14	8	1.7	236	UBIG_HABDU	Q7vkw2 haemophilus
15	8	1.7	236	UBIG_SHEON	Q8ee99 shewanella
16	8	1.7	240	UBIG_ECOL6	Q8xe29 escherichia
17	8	1.7	240	UBIG_ECOL6	Q8ffp0 escherichia
18	8	1.7	240	UBIG_ECOL1	P17993 escherichia
19	8	1.7	240	UBIG_PHOLL	Q7n2m5 photorhabdu
20	8	1.7	240	UBIG_SHIFL	Q820c5 shigella fl
21	8	1.7	242	UBIG_PASMU	Q9cm16 pasteurella
22	8	1.7	242	UBIG_SALTU	Q8z560 salmonella
23	8	1.7	242	UBIG_SALTU	P37431 salmonella
24	8	1.7	242	UBIG_YERPE	Q8z9re yersinia pe
25	8	1.7	342	1RMC_ECOLI	P39406 escherichia
26	8	1.7	348	1GTM_ARATH	Q9zak1 arabidopsis
27	8	1.7	376	1ERG_CANAL	Q14198 candida alb
28	8	1.7	501	1TRA2_HUMAN	Q12933 homo sapien
29	8	1.7	501	1TRA2_MOUSE	P39429 mus musculu
30	8	1.7	572	1GUXA_CELFI	P50401 cellulomona
31	7	1.5	71	1YQOQ_BACSU	P54494 bacillus su
32	7	1.5	111	1SINR_BACLI	P22753 bacillus li
33	7	1.5	111	1SINR_BACSU	P06533 bacillus su

## RESULT 1

ID	PEAM_SPIOL	STANDARD;	PRT,	494 AA.
AC	Q9M571;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Phosphoethanolamine N-methyltransferase (EC 2.1.1.103).			
GN	PEAMT.			
OS	Spinacia oleracea (Spinach).			
OC	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
OC	Caryophyllales; Amaranthaceae; Spinacia.			
OX	NCBI_TaxID=3562;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND CHARACTERIZATION.			
RC	STRAIN=cv. Savoy hybrid 612; TISSUE=Leaf.			
RX	MEDLINE=20261526; PubMed=10799484;			
RA	Nuccio M.L., Ziemak M.J., Henry S.A., Weretilnyk E.A., Hanson A.D.;			
RT	"Phosphoethanolamine N-methyltransferase from spinach: cDNA cloning by			
RT	complementation in Schizosaccharomyces pombe and characterization of			
RT	the recombinant enzyme."			
RL	J. Biol. Chem. 275:14095-14101(2000).			
CC	FUNCTION: Catalyzes N-methylation of phosphoethanolamine,			
CC	phosphomonomethyl ethanolamine and phosphodimethyl ethanolamine, the			
CC	three methylation steps required to convert phosphoethanolamine to			
CC	phosphocholine. Mediates a key step in the biosynthesis of			
CC	choline, a precursor of the osmoprotectant glycine betaine. Has an			
CC	optimal pH of 7.8 to 8.5. Has no ethanolamine- or			
CC	phosphatidylethanolamine-N-methyltransferase activity.			
CC	CATALYTIC ACTIVITY: S-adenosyl-L-methionine + ethanolamine			
CC	phosphate = S-adenosyl-L-homocysteine + N-methylethanolamine			
CC	phosphate.			
CC	ENZYME REGULATION: Inhibited by phosphocholine but not by choline,			
CC	glycine betaine, monomethyl ethanolamine or dimethylethanolamine.			
CC	SUBUNIT: Monomer.			
CC	SUBCELLULAR LOCATION: Cytoplasmic.			
CC	INDUCTION: By salt stress.			
CC	SIMILARITY: Belongs to the methyltransferase superfamily.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	EMBL; AF237623; AAP61950.1;			
DR	InterPro; IPR001601; Methyltransf.			
DR	InterPro; IPR000051; SAM_bind.			
DR	Methyltransferase; Transferase; Repeat.			
KM	DOMAIN 53...162 ...SAM-BINDING 1.			
FT	DOMAIN 282...389 SAM-BINDING 2.			
FT	SEQUENCE 494 AA; 56361 MW; 7F2537C8E4B8413B CRC64;			

P00217 haloarcula  
P00216 halobacteri  
P25649 saccharomyc  
O67445 aquifex aeo  
Q8rdm9 fusbacteri  
Q10886 mycobacteri  
Q9c3n8 pasteurella  
P77889 lactobacill  
Q35423 synechocyst  
Q9cbb8 mycobacteri  
O06424 mycobacteri  
Q82055 coxiella bu

## ALIGNMENTS

Query Match 9.5%; Score 45; DB 1; Length 494;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-37;  
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC 122 LVERMLKWKCGVYFRESCHQSDHKKSNPTHYRPRVTK 166  
 DB 143 LVERMLKWKCGVYFRESCHQSDHKKSNPTHYRPRVTK 187

RESULT 2

PEM2 ARATH  
 ID PEM2 ARATH STANDARD; PRT; 475 AA.  
 AC Q944H0; Q9LP63; Q9LP64;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Putative phosphoethanolamine N-methyltransferase 2 (EC 2.1.1.103).  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 CC eurosid II; Brassicales; Brassicaceae; Arabidopsids.  
 CC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=21016719; PubMed=11130712;  
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,  
 RA White O., Alonso J., Altai H., Araujo R., Bowman C.L., Brooks S.Y.,  
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,  
 RA Chung M.K., Conn L., Conway A.B., Cressey T.H., Dewar K.,  
 RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
 RA Gill J.B., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,  
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,  
 RA Langin-Hooper S., Lee A., Lee J.M., Luros J.S., Maiti R., Marziani A.,  
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,  
 RA Miltischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,  
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
 RA Sakano H., Salzman S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
 RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,  
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
 RA "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
 thaliana";  
 RL Nature 408:816-820(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=22954850; PubMed=14593172;  
 RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,  
 RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,  
 RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,  
 RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,  
 RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,  
 RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,  
 RA Chao Q., Choy N., Egu A., Goldsmith A.D., Gurjal M., Hansen N.F.,  
 RA Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,  
 RA Kham S., Koesema G., Ishida J., Jiang P.X., Jones T., Kawai J.,  
 RA Kaniya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,  
 RA Satou M., Tamari E., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,  
 RA Tuan S., Shinohara K., Davis R.W., Theologis A., Ecker J.R.;  
 RA "Empirical analysis of transcriptional activity in the Arabidopsis  
 genome";  
 RL Science 302:842-846(2003).  
 CC -1- FUNCTION: Catalyzes N-methylation of phosphoethanolamine,  
 CC phosphomonomethyl ethanolamine and phosphodimethyl ethanolamine, the  
 CC three methylation steps required to convert phosphoethanolamine to  
 CC phosphocholine (By similarity).  
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + ethanolamine  
 CC phosphate = S-adenosyl-L-homocysteine + N-methylethanolamine  
 CC phosphate.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -1- SIMILARITY: Belongs to the methyltransferase superfamily.  
 CC -1- CAUTION: Ref.1 sequences differ from that shown due to erroneous  
 CC gene model prediction.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL; AC020889; AAF79704.1; ALT\_SEQ.  
 CC EMBL; AF428454; ALJ16223.1;  
 CC EMBL; AY063866; ALJ36222.1;  
 CC InterPro; IPR001601; Methyltransf.  
 CC InterPro; IPR000051; SAM\_Bind.  
 CC InterPro; IPR004033; UbiE/COOS\_Metrf.  
 CC Pfam; PF01209; UbiE\_methyltran; 1.  
 CC XW Transferase; Methyltransferase; Repeat.  
 CC FT DOMAIN 34 143  
 CC SAM-BINDING 1.  
 CC FT DOMAIN 263 370  
 CC SAM-BINDING 2.  
 CC SQ SEQUENCE 475 AA; 54018 MW; 040705A4ADE80DA8 CRC64;

Query Match 5.5%; Score 26; DB 1; Length 475;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-18;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 YPDNSFDVYSRDTILHIQDKPALFR 351  
 DB 328 YPDNSFDVYSRDTILHIQDKPALFR 353

RESULT 3

PEM3 ARATH  
 ID PEM3 ARATH STANDARD; PRT; 490 AA.  
 AC Q9CS99; Q9C9V1;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Putative phosphoethanolamine N-methyltransferase 3 (EC 2.1.1.103).  
 GN NMT3 OR AT1G73600 OR F6D5.1 OR F25P22.1.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 CC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 CC eurosid II; Brassicales; Brassicaceae; Arabidopsids.  
 CC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=21016719; PubMed=11130712;  
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,  
 RA White O., Alonso J., Altai H., Araujo R., Bowman C.L., Brooks S.Y.,  
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,  
 RA Chung M.K., Conn L., Conway A.B., Cressey T.H., Dewar K.,  
 RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
 RA Gill J.B., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,  
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,  
 RA Langin-Hooper S., Lee A., Lee J.M., Luros J.S., Maiti R., Marziani A.,  
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,  
 RA Miltischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,  
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
 RA Sakano H., Salzman S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
 RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,  
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
 RA "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
 thaliana";  
 RL Nature 408:816-820(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=22954850; PubMed=14593172;  
 RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,  
 RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,  
 RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,  
 RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,  
 RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,  
 RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,  
 RA Chao Q., Choy N., Egu A., Goldsmith A.D., Gurjal M., Hansen N.F.,  
 RA Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,  
 RA Kham S., Koesema G., Ishida J., Jiang P.X., Jones T., Kawai J.,  
 RA Kaniya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,  
 RA Satou M., Tamari E., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,  
 RA Tuan S., Shinohara K., Davis R.W., Theologis A., Ecker J.R.;  
 RA "Empirical analysis of transcriptional activity in the Arabidopsis  
 genome";  
 RL Science 302:842-846(2003).  
 CC -1- FUNCTION: Catalyzes N-methylation of phosphoethanolamine,  
 CC phosphomonomethyl ethanolamine and phosphodimethyl ethanolamine, the  
 CC three methylation steps required to convert phosphoethanolamine to  
 CC phosphocholine (By similarity).  
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + ethanolamine  
 CC phosphate = S-adenosyl-L-homocysteine + N-methylethanolamine  
 CC phosphate.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).



OS Methanopyrus kandleri.  
 OC Archaea: Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;  
 OC Methanopyrus.  
 OX NCBI\_TaxID=2320;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AV19 / DSM 6324 / JCM 9639;  
 RX MEDLINE=21927647; PubMed=11930014;  
 RA Shestakov A.I., Mezheva V.V., Makarova K.S., Polushin N.N.,  
 RA Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,  
 RA Nale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,  
 RA Malykh A.G., Kozlov E.V., Koryavkin S.A.,  
 RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19  
 RT and monophyly of archaeal methanogens";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649 (2002).  
 CC -1- FUNCTION: Catalyzes the 1,3-allylic rearrangement of the  
 CC homoallylic substrate isopentenyl pyrophosphate (IPP) to its allylic isomer,  
 CC dimethylallyl pyrophosphate (DMAPP) (By similarity).  
 CC -1- CATALYTIC ACTIVITY: isopentenyl diphosphate = dimethylallyl  
 CC diphosphate.  
 CC -1- COFACTOR: FMN and NADPH (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: Belongs to the IPP isomerase type 2 family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AF010369; AAM01990.1; ALT\_INIT.  
 CC HAMAP; MF\_00354; -; 1.  
 CC InterPro; IPR003009; FMN enzyme.  
 CC InterPro; IPR00262; FMN\_hydryac\_dh.  
 CC Pfam; PF01070; FMN dh; 1.  
 CC Isomerase; isoprene biosynthesis; Flavoprotein; FMN; NADP;  
 CC Complete proteome.  
 CC SEQUENCE 365 AA; 39272 MW; B5C0541EBB49C355 CRC64;  
 SQ  
 Query Match 1.9%; Score 9; DB 1; Length 365;  
 Best Local Similarity 100.0%; Pred. No. 0.54;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 403 AGFVDVLAE 411  
 DB 167 AGFVDVLAE 175  
 ID GLCF\_ECOLI  
 AC P52074; P76654; STANDARD; PRT; 407 AA.  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Glycolate oxidase iron-sulfur subunit.  
 GN GLCF OR GOX B2978.  
 OS Escherichia coli  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / W3110;  
 RX MEDLINE=96178972; PubMed=8606183;  
 RA Pellicer M.T., Badia J., Aguilar J.T., Baldoma L.,  
 RA "The glc locus of Escherichia coli: characterization of genes encoding  
 RT the subunits of glycolate oxidase and the glc regulator protein.";  
 RL J. Bacteriol. 178:2051-2059 (1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.

RC STRAIN=K12 / M31655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blactner F.R., Plunkett G., III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glaeser J.D., Rode C.K., Mayhew G.P.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.,  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:1232-1244 (1997).  
 CC -1- SIMILARITY: The iron-sulfur centers are similar to those of  
 CC bacterial-type 4Fe-4S ferredoxins.  
 CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A  
 CC FRAMESHIFT THAT FUSES TOGETHER GLCE AND GLCF.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; L43490; AAB02532.1; -;  
 CC EMBL; U28377; AAB69145.1; ALT\_FRAME.  
 CC EMBL; AE000380; AAC76014.1; ALT\_FRAME.  
 CC EcoGene; EG13291; glcf.  
 CC InterPro; IPR001450; 4Fe4S ferredoxin.  
 CC InterPro; IPR004017; DUF224.  
 CC Pfam; PF02754; CCG; 2.  
 CC Pfam; PF00037; fer4; 2.  
 CC PROSITE; PS00198; 4Fe4S\_FERREDOXIN; 2.  
 CC Electron transport; Iron-sulfur; 4Fe-4S; Complete proteome.  
 CC METAL 25 25 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).  
 CC FT METAL 28 28 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).  
 CC FT METAL 31 31 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).  
 CC FT METAL 35 35 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).  
 CC FT METAL 75 75 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).  
 CC FT METAL 78 78 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).  
 CC FT METAL 81 81 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).  
 CC FT METAL 85 85 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).  
 CC FT METAL 85 85 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).  
 CC SEQUENCE 407 AA; 45110 MW; F983C893063F76DC CRC64;  
 SQ  
 Query Match 1.9%; Score 9; DB 1; Length 407;  
 Best Local Similarity 100.0%; Pred. No. 0.59;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 393 VKKEYQMLK 401  
 DB 255 VKKEYQMLK 263  
 ID CYAA\_SACKL  
 AC P23466; STANDARD; PRT; 1839 AA.  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Adenylate cyclase (EC 4.6.1.1) (ATP pyrophosphate-lyase) (Adenylyl  
 DE cyclase).  
 GN CYR.  
 OS Saccharomyces kluyveri (Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4934;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9132718; PubMed=1864503;  
 RA Young D., O'Neill K., Broek D., Wigler M.,  
 RA "The adenylate cyclase-encoding gene from Saccharomyces kluyveri.";  
 RL Gene 102:129-132 (1991).  
 CC -1- FUNCTION: Plays essential roles in regulation of cellular  
 CC metabolism by catalyzing the synthesis of a second messenger,  
 CC cAMP.





UBIG\_PSBPK STANDARD; PRT; 232 AA.

AC Q88M10;  
 DT 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE 3-demethylubiquinone-9 3-methyltransferase (SC 2.1.1.64) (3,4-  
 dihydroxy-5-hexaprenylbenzoate methyltransferase) (DHHE  
 methyltransferase)  
 DE UBIG OR PFI765.  
 GN Pseudomonas putida (strain KT2440).  
 OS Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=160488;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22423060; PubMed=12534463;  
 RA Nelson K.E., Weisel C., Paulsen I.T., Dodson R.J., Hilbert H.,  
 Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,  
 Brinkac L., Beanan M.W., Deboy R.T., Daugherty S., Kolonay J.,  
 Madupu R., Nelson M., White O., Peterson J., Khouri H., Hance I.,  
 Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzes A.,  
 Utterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,  
 Lauber J., Stjepandic D., Hohnselt J., Straetz M., Helm S.,  
 Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,  
 Fraser C.M.;  
 RA "Complete genome sequence and comparative analysis of the  
 RT metabolically versatile Pseudomonas putida KT2440.";  
 RL Environ. Microbiol. 4:799-808(2002).  
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3-  
 CC demethylubiquinone-9 = S-adenosyl-L-homocysteine + ubiquinone-9.  
 CC -1- PATHWAY: Ubiquinone biosynthesis.  
 CC -1- SIMILARITY: Belongs to the ubiG/COQ3 family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AS016780; AAN67385.1; -  
 CC TIGR; PFI765; -  
 CC HAMAP; MF\_00472; -; 1.  
 CC InterPro; IPR001601; Methyltransf.  
 CC InterPro; IPR000051; SAM bind.  
 CC InterPro; IPR004033; UbiE/COQ5 Metrf.  
 CC Pfam; PF01209; Ubie\_methyltran; 1.  
 CC TIGRFAMs; TIGR01983; UbiG; 1.  
 CC Ubiquinone biosynthesis; Transferase; Methyltransferase;  
 CC Complete proteome.  
 CC SEQUENCE 232 AA; 26084 MW; 7A795B7DF50479B9 CRC64;  
 Query Match 1.7%; Score 8; DB 1; Length 232;  
 Best Local Similarity 100.0%; Pred. No. 3.9;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 267 KVLVDGCG 274  
 DB 50 KVLVDGCG 57  
 RESULT 10  
 UBIG\_PSESM STANDARD; PRT; 232 AA.  
 AC Q885T9;  
 DT 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE 3-demethylubiquinone-9 3-methyltransferase (SC 2.1.1.64) (3,4-  
 dihydroxy-5-hexaprenylbenzoate methyltransferase) (DHHE  
 methyltransferase).  
 DE UBIG OR PFI765.  
 GN Pseudomonas putida (strain KT2440).  
 OS Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=160488;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22423060; PubMed=12534463;  
 RA Nelson K.E., Weisel C., Paulsen I.T., Dodson R.J., Hilbert H.,  
 Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,  
 Brinkac L., Beanan M.W., Deboy R.T., Daugherty S., Kolonay J.,  
 Madupu R., Nelson M., White O., Peterson J., Khouri H., Hance I.,  
 Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzes A.,  
 Utterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,  
 Lauber J., Stjepandic D., Hohnselt J., Straetz M., Helm S.,  
 Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,  
 Fraser C.M.;  
 RA "Complete genome sequence and comparative analysis of the  
 RT metabolically versatile Pseudomonas putida KT2440.";  
 RL Environ. Microbiol. 4:799-808(2002).  
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3-  
 CC demethylubiquinone-9 = S-adenosyl-L-homocysteine + ubiquinone-9.  
 CC -1- PATHWAY: Ubiquinone biosynthesis.  
 CC -1- SIMILARITY: Belongs to the ubiG/COQ3 family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AS016780; AAN67385.1; -  
 CC TIGR; PFI765; -  
 CC HAMAP; MF\_00472; -; 1.  
 CC InterPro; IPR001601; Methyltransf.  
 CC InterPro; IPR000051; SAM bind.  
 CC InterPro; IPR004033; UbiE/COQ5 Metrf.  
 CC Pfam; PF01209; Ubie\_methyltran; 1.  
 CC TIGRFAMs; TIGR01983; UbiG; 1.  
 CC Ubiquinone biosynthesis; Transferase; Methyltransferase;  
 CC Complete proteome.  
 CC SEQUENCE 232 AA; 26084 MW; 7A795B7DF50479B9 CRC64;  
 Query Match 1.7%; Score 8; DB 1; Length 232;  
 Best Local Similarity 100.0%; Pred. No. 3.9;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 267 KVLVDGCG 274  
 DB 50 KVLVDGCG 57

UBIG OR PSPT01742.  
 OS Pseudomonas syringae (pv. tomato).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=323;  
 RN (1)\_TaxID=323;  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=DC3000;  
 RA MEDLINE=22834015; PubMed=12928499;  
 RA Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,  
 Quinn M.L., Dodson R.J., Deboy R.T., Durkin A.S., Kolonay J.F.,  
 Madupu R., Daugherty S., Brinkac L., Beanan M.J., Haft D.H.,  
 Nelson W.C., Davidson T., Zafar N., Zhou L., Liu J., Yuan Q.,  
 Khouri H., Fedorova N., Tran B., Russell D., Berry K., Utterback T.,  
 Van Aken S.E., Feldblum T.V., D'Ascenzo M., Deng W.-L., Ramos A.R.,  
 Alfano J.R., Cartinhour S., Chatterjee A.K., Delaney T.P.,  
 Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X., Bender C.L.,  
 White G., Fraser C.M., Collier A.;  
 RA "The complete genome sequence of the Arabidopsis and tomato pathogen  
 RT Pseudomonas syringae pv. tomato DC3000.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).  
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3-  
 CC demethylubiquinone-9 = S-adenosyl-L-homocysteine + ubiquinone-9.  
 CC -1- PATHWAY: Ubiquinone biosynthesis.  
 CC -1- SIMILARITY: Belongs to the ubiG/COQ3 family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AE016862; AOS55262.1; ALT\_INIT.  
 CC TIGR; PSPT01742; -; 1.  
 CC HAMAP; MF\_00472; -; 1.  
 CC InterPro; IPR001601; Methyltransf.  
 CC InterPro; IPR000051; SAM bind.  
 CC InterPro; IPR004033; UbiE/COQ5 Metrf.  
 CC Pfam; PF01209; Ubie\_methyltran; 1.  
 CC TIGRFAMs; TIGR01983; UbiG; 1.  
 CC Ubiquinone biosynthesis; Transferase; Methyltransferase;  
 CC Complete proteome.  
 CC SEQUENCE 232 AA; 25932 MW; 248F3D231401A9E4 CRC64;  
 Query Match 1.7%; Score 8; DB 1; Length 232;  
 Best Local Similarity 100.0%; Pred. No. 3.9;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 267 KVLVDGCG 274  
 DB 50 KVLVDGCG 57  
 RESULT 11  
 UBIG\_VIBPA STANDARD; PRT; 235 AA.  
 AC Q87ND5;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE 3-demethylubiquinone-9 3-methyltransferase (SC 2.1.1.64) (3,4-  
 dihydroxy-5-hexaprenylbenzoate methyltransferase) (DHHE  
 methyltransferase).  
 DE UBIG OR VPI933.  
 GN Vibrio parahaemolyticus.  
 OS Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
 OC Vibrionaceae; Vibrio.  
 OX NCBI\_TaxID=670;  
 RN (1)\_TaxID=670;  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=RIMD 2210633 / Serotype O3:K6;  
 RC



```

RX MEDLINE=22509454; PubMed=12620739;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yaenaga T., Honda T., Shinagawa H., Hattori M., Iida T.,
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae."
RL Lancet 361:743-749(2003).
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3-
CC demethylubiquinone-9 = S-adenosyl-L-homocysteine + ubiquinone-9.
CC -!- PATHWAY: Ubiquinone biosynthesis.
CC -!- SIMILARITY: Belongs to the ubiG/COQ3 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP005079; BAC60136.1; -.
DR HAMAP; MF_00472; -.
DR InterPro; IPR001601; Methyltransf.
DR UbiG; UBI000051; SAM_bind.
KW Ubiquinone biosynthesis; Transferase; Methyltransferase;
KW Complete proteome.
SQ SEQUENCE 235 AA; 26179 MW; 934D8E4403D017F0 CRC64;

Query Match 1.7%; Score 8; DB 1; Length 235;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 KVLVDGCG 274
DB 54 KVLVDGCG 61

RESULT 12
UBIG_VIBVU STANDARD; PRT; 235 AA.
AC QD8E9; (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 3-demethylubiquinone-9 3-methyltransferase (EC 2.1.1.64) (3,4-
DE dihydroxy-5-hexaprenylbenzoate methyltransferase) (DHBB
DE methyltransferase).
DE UBIG OR VV13040.
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMC6."
RT Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3-
CC demethylubiquinone-9 = S-adenosyl-L-homocysteine + ubiquinone-9.
CC -!- PATHWAY: Ubiquinone biosynthesis.
CC -!- SIMILARITY: Belongs to the ubiG/COQ3 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE016807; AN011364.1; -.

```

```

DR HAMAP; MF_00472; -.
DR InterPro; IPR001601; Methyltransf.
DR UbiG; UBI000051; SAM_bind.
KW Ubiquinone biosynthesis; Transferase; Methyltransferase;
KW Complete proteome.
SQ SEQUENCE 235 AA; 26244 MW; A89B7E1F8613F2B8 CRC64;

Query Match 1.7%; Score 8; DB 1; Length 235;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 KVLVDGCG 274
DB 54 KVLVDGCG 61

RESULT 13
UBIG_VIBVU STANDARD; PRT; 235 AA.
AC Q7NM27;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE 3-demethylubiquinone-9 3-methyltransferase (EC 2.1.1.64) (3,4-
DE dihydroxy-5-hexaprenylbenzoate methyltransferase) (DHBB
DE methyltransferase).
DE UBIG OR VV1246.
OS Vibrio vulnificus (strain YJ016).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=196600;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen C.Y., Wu K.M., Chang Y.C., Chang C.H., Tsai H.C., Liao T.L.,
RA Liu Y.M., Chen H.J., Shen A.B., Li J.C., Su T.L., Shao C.P., Lee C.T.,
RA Hor L.I., Tsai S.F.;
RT "Comparative genome analysis of Vibrio vulnificus, a marine
RT pathogen."
RT Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3-
CC demethylubiquinone-9 = S-adenosyl-L-homocysteine + ubiquinone-9.
CC -!- PATHWAY: Ubiquinone biosynthesis.
CC -!- SIMILARITY: Belongs to the ubiG/COQ3 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP005335; BAC94010.1; -.
DR HAMAP; MF_00472; -.
DR UbiG; UBI000051; SAM_bind.
KW Ubiquinone biosynthesis; Transferase; Methyltransferase;
KW Complete proteome.
SQ SEQUENCE 235 AA; 26244 MW; A89B7E1F8613F2B8 CRC64;

Query Match 1.7%; Score 8; DB 1; Length 235;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 KVLVDGCG 274
DB 54 KVLVDGCG 61

RESULT 14
UBIG_HABDU STANDARD; PRT; 236 AA.
AC Q7VKW2;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)

```

DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE 3-demethylubiquinone-9 3-methyltransferase (EC 2.1.1.64) (3,4-  
 DE dihydroxy-5-hexaprenylbenzoate methyltransferase) (DHHS  
 DE methyltransferase)  
 GN UBIG OR HD1750.  
 OS Haemophilus ducreyi.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
 OC Pasteurellaceae; Haemophilus.  
 OX NCBI\_TaxID=730;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=3500HP / ATCC 700724;  
 RA Munson R.S. Jr., Ray W.C., Mahairas G., Sabo P., Mungur R.,  
 RA Johnson L., Nguyen D., Wang J., Forst C., Hood L.;  
 RA "The complete genome sequence of Haemophilus ducreyi.";  
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3-  
 CC demethylubiquinone-9 = S-adenosyl-L-homocysteine + ubiquinone-9.  
 CC -1- PATHWAY: Ubiquinone biosynthesis.  
 CC -1- SIMILARITY: Belongs to the ubiG/COQ3 family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; AB017155; AAP96505.1; -;  
 DR HAMAP; MF 00472; -; 1.  
 DR InterPro; IPR001601; Methyltransf.  
 DR InterPro; IPR000051; SAM bind.  
 DR TIGRFAMs; TIGR01983; UbiG; 1.  
 DR UbiG; 1.  
 KW Ubiquinone biosynthesis; Transferase; Methyltransferase;  
 KW Complete proteome.  
 SQ SEQUENCE 236 AA; 26602 MW; B266FC407FF6AECDC CRC64;  
 Query Match 1.7%; Score 8; DB 1; Length 236;  
 Best Local Similarity 100.0%; Pred. No. 3.9;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 267 KVLVDGCG 274  
 DB 51 KVLVDGCG 58  
 RESULT 15  
 UBIG SHEON  
 ID UBIG SHEON STANDARD; PRT; 236 AA.  
 AC Q8EG9;  
 DT 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE 3-demethylubiquinone-9 3-methyltransferase (EC 2.1.1.64) (3,4-  
 DE dihydroxy-5-hexaprenylbenzoate methyltransferase) (DHHS  
 DE methyltransferase)  
 GN UBIG OR SO2413.  
 OS Shewanella oneidensis.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;  
 OC Alteromonadaceae; Shewanella.  
 OX NCBI\_TaxID=70863;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NR-1;  
 RX MEDLINE=22297686; PubMed=12368813;  
 RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,  
 RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,  
 RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,  
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,  
 RA Maduppi K., Peterson J.D., Umayam L.A., White O., Wolf A.M.,  
 RA Vamathevan J., Weidman J., Imbraim M., Lee K., Berry K., Lee C.,  
 RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,

RA Felblyum T.V., Smith H.O., Venter J.C., Nealeon K.H., Fraser C.M.;  
 RT "Genome sequence of the dissimilatory metal ion-reducing bacterium  
 RL *Shewanella oneidensis*.";  
 RL Nat. Biotechnol. 20:1118-1123(2002).  
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3-  
 CC demethylubiquinone-9 = S-adenosyl-L-homocysteine + ubiquinone-9.  
 CC -1- PATHWAY: Ubiquinone biosynthesis.  
 CC -1- SIMILARITY: Belongs to the ubiG/COQ3 family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; AB015682; AAN55447.1; -;  
 DR TIGR; SO2413; -; 1.  
 DR HAMAP; MF 00472; -; 1.  
 DR InterPro; IPR001601; Methyltransf.  
 DR InterPro; IPR000051; SAM bind.  
 DR TIGRFAMs; TIGR01983; UbiG; 1.  
 DR UbiG; 1.  
 KW Ubiquinone biosynthesis; Transferase; Methyltransferase;  
 KW Complete proteome.  
 SQ SEQUENCE 236 AA; 26469 MW; A4F0C932B5533E76 CRC64;  
 Query Match 1.7%; Score 8; DB 1; Length 236;  
 Best Local Similarity 100.0%; Pred. No. 3.9;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 267 KVLVDGCG 274  
 DB 54 KVLVDGCG 61  
 Search completed: July 26, 2004, 13:22:37  
 Job time : 15 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2004, 13:19:37 ; Search time 42 Seconds  
(Without alignments)  
3553.336 Million cell updates/sec

Title: US-10-031-331B-40  
Perfect score: 473  
Sequence: 1 HTVDLTTEAMMLDSQASDLDD.....KLRRTAKGEQWGLFVAXKK 473

Scoring table: OUIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1017041 seqs, 315518202 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SPTREMBL 25:\*

- 1: sp\_archea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	473	100.0	494	10 Q852S7	Q852S7 suaeda japo
2	27	5.7	491	10 Q9AXH3	Q9AXH3 lycopersico
3	26	5.5	376	10 Q8L7A8	Q8L7A8 arabidopsis
4	26	5.5	493	10 Q846A4	Q846A4 aster tripo
5	24	5.1	491	10 Q7XJ22	Q7XJ22 brassica na
6	21	4.4	498	10 Q8VYX1	Q8VYX1 triticum ae
7	17	3.6	499	10 Q8LJ10	Q8LJ10 oryza sativ
8	16	3.4	346	10 Q82720	Q82720 nicotiana t
9	16	3.4	349	10 Q82434	Q82434 nicotiana t
10	16	3.4	367	10 Q43445	Q43445 glycine max
11	14	3.0	336	10 Q8LKW1	Q8LKW1 arabidopsis
12	14	3.0	336	10 Q9LM02	Q9LM02 arabidopsis
13	12	2.5	346	10 Q24328	Q24328 ricinus com
14	11	2.3	330	10 Q84M50	Q84M50 oryza sativ
15	11	2.3	344	10 P93852	P93852 zea mays (m
16	11	2.3	344	10 Q49215	Q49215 zea mays (m

17	11	2.3	349	10 Q82426	Q82426 oryza sativ
18	11	2.3	363	10 Q41587	Q41587 triticum ae
19	11	2.3	363	10 Q41586	Q41586 triticum ae
20	10	2.1	87	16 Q7VD26	Q7VD26 prochloroco
21	10	2.1	304	16 Q7V427	Q7V427 prochloroco
22	10	2.1	309	16 Q7U4C8	Q7U4C8 synchococ
23	10	2.1	318	16 P74388	P74388 synchococ
24	10	2.1	328	16 Q8DI65	Q8DI65 synchococ
25	9	1.9	247	16 Q982P3	Q982P3 rhiobium l
26	9	1.9	255	16 Q81TP8	Q81TP8 bacillus an
27	9	1.9	273	2 Q8XI52	Q8XI52 nocardia ae
28	9	1.9	280	16 Q8YWL3	Q8YWL3 anabaena sp
29	9	1.9	283	2 Q8KZ94	Q8KZ94 nocardia ae
30	9	1.9	311	16 Q7UZZ0	Q7UZZ0 prochloroco
31	9	1.9	330	16 Q8YV60	Q8YV60 anabaena sp
32	9	1.9	341	5 Q86I59	Q86I59 dictyosteli
33	9	1.9	437	5 Q22993	Q22993 caenorhabdi
34	9	1.9	2094	11 Q80Y35	Q80Y35 mus musculu
35	8	1.7	88	17 Q970B1	Q970B1 sulfolobus
36	8	1.7	202	17 Q8PU17	Q8PU17 methanosarc
37	8	1.7	207	16 Q8XNS1	Q8XNS1 clostridium
38	8	1.7	221	16 Q8P5S7	Q8P5S7 leptospira
39	8	1.7	223	16 Q89UT6	Q89UT6 bradyrhizob
40	8	1.7	228	16 P96366	P96366 mycobacteri
41	8	1.7	228	16 Q7U0U7	Q7U0U7 mycobacteri
42	8	1.7	229	17 Q8TRV4	Q8TRV4 methanosarc
43	8	1.7	232	16 Q88M10	Q88M10 pseudomonas
44	8	1.7	236	16 Q8EEG9	Q8EEG9 shewanella
45	8	1.7	236	16 Q7VKW2	Q7VKW2 haemophilus

## ALIGNMENTS

## RESULT 1

Q852S7 PRELIMINARY; PRT; 494 AA.

AC Q852S7; DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Phosphoethanolamine N-methyltransferase.

GN PEAUT.

OS Sueda japonica.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Caryophyllales; Amaranthaceae; Sueda.

OX NCBI\_TaxID=90346;

RN [1]

RP SEQUENCE FROM N.A.

RA Yamada A., Nozawa G.T., Tanimoto S., Ozeki Y.;

RT "Glycinebetaine synthesis in Sueda japonica.";

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB080186; BAC57432.1; "

DR GO; GO:0008757; F:8-adenosylmethionine-dependent methyltransf. . . ; IEA.

DR GO; GO:0016740; F:transferase activity; IEA.

DR InterPro; IPR001601; Methyltransf.

DR InterPro; IPR000051; SAM\_Bind.

DR InterPro; IPR004033; UBIIE/COO5\_Metrf.

DR Pfam; PF01209; Ubie methyltran; 1.

KW Transferase; Methyltransferase.

SQ SEQUENCE 494 AA; 56562 MW; BC613P9097BD3AE3 CRC64;

Query Match 100.0%; Score 473; DB 10; Length 494;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 473; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HTVDLTTEAMMLDSQASDLDDKERPERILSNLPPLPGLCKLLELGACIGFTGELAFKAGQV 60

DB 27 HTVDLTTEAMMLDSQASDLDDKERPERILSNLPPLPGLCKLLELGACIGFTGELAFKAGQV 81

QY 61 TALDPTESAIKKKENVINGHYKNVKNPCADVTSPTLSFPFHSIDLVIFSNWLLWYLSDEVE 120

Db 82 IALDPIESAIKNEVINGHYKNKWCADVTSTLSPFPHSLDVFISNMLLMVLSDEVE 141  
QY 121 NLVERMLKMLKPGYIFPESCHQSDHRSNTHVREPRFYTKAFKCHLODGSNS 180  
Db 142 NLVERMLKMLKPGYIFPESCHQSDHRSNTHVREPRFYTKAFKCHLODGSNS 201  
QY 181 YELSLSCKICIGAVVRNKNQNIQISWLMKQVSKDDKQFRLDTSQYKCNILRYERV 240  
Db 202 YELSLSCKICIGAVVRNKNQNIQISWLMKQVSKDDKQFRLDTSQYKCNILRYERV 261  
QY 241 GPGVSTGGVETTKFVSMMLDKPKQKVLVDGCGIGGDFYMAETFDVEVVGFLSVNMI 300  
Db 262 GPGVSTGGVETTKFVSMMLDKPKQKVLVDGCGIGGDFYMAETFDVEVVGFLSVNMI 321  
QY 301 SPALERSLGLKCAVEFVADCTKINPDNSFDVYSRDTILHIQDKPALFRSFKMLKPG 360  
Db 322 SPALERSLGLKCAVEFVADCTKINPDNSFDVYSRDTILHIQDKPALFRSFKMLKPG 381  
QY 361 GRVLISYCKKAGPPSPFAAYTKQGYDLHVKYQMLKQDAGFVVLADRTQFIRV 420  
Db 382 GRVLISYCKKAGPPSPFAAYTKQGYDLHVKYQMLKQDAGFVVLADRTQFIRV 441  
QY 421 LRKLETVKEKDVIFISDFSEEDYNDIVGQNDKLRATKAGQRWGLFVAKK 473  
Db 442 LRKLETVKEKDVIFISDFSEEDYNDIVGQNDKLRATKAGQRWGLFVAKK 494

RESULT 2  
Q9AXH3 PRELIMINARY; PRT; 491 AA.  
AC Q9AXH3; 17, Created)  
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)  
DT 01-JUN-2001 (TRENBLrel. 24, Last annotation update)  
DE Phosphoethanolamine N-methyltransferase.  
OS Lycopersicon esculentum (Tomato).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC Lamiales; Solanales; Solanaceae; Solanum.  
OX NCBI\_TaxID=4081;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Parani M., Parida A.;  
RT "Characterization of a cDNA for phosphoethanolamine N-methyltransferase."  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF328858; AGS9894.1; --  
DR GO; GO:0005344; P:cytochrome transporter activity; IEA.  
DR GO; GO:0008757; P:S-adenosylmethionine-dependent methyltransferase.  
DR GO; GO:0016740; P:transferase activity; IEA.  
DR GO; GO:0006810; P:transferase; IEA.  
DR InterPro; IPR000896; Hemocyanin.  
DR InterPro; IPR001601; Methyltransf.  
DR InterPro; IPR000051; SAM bind.  
DR PROSITE; PS00210; HEMOCYANIN.2.1.  
DR Methyltransferase, Transferase.  
SQ SEQUENCE 491 AA; 55947 MW; D36RCODBS12733E7 CRC64;

Query Match 5.7%; Score 27; DB 10; Length 491;  
Best Local Similarity 100.0%; Pred. No. 4.8e-19;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 331 FVIVSRDTILHIQDKPALFRSFKYWL 357  
Db 349 FVIVSRDTILHIQDKPALFRSFKYWL 375

RESULT 3  
Q8L7A8 PRELIMINARY; PRT; 376 AA.  
AC Q8L7A8; 22, Created)  
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)  
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)

DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
DE Phosphoethanolamine N-methyltransferase, putative.  
GN ATIG73600.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC Eurosidia II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Southwick P., Nguyen M., Tripp M., Palm C.J., Jones T., Wu T.,  
RA Carninci P., Chen H., Cheuk R., Chan M.M., Chang C.H., Dale J.M.,  
RA Deng J.M., Hayashizaki Y., Heuan J.M., Lee J.M., Ishida J., Kamiya A.,  
RA Kawai J., Kim C.J., Narusaka M., Quach H.L., Sakurai T., Satou M.,  
RA Seki M., Shinn P., Tang C.C., Toroumi M., Wallender B.K., Wong C.,  
RA Wu H.C., Yamada K., Yu G., Yuan S., Shinzaki K., Ecker J.,  
RA Theologis A., Davis R.W.;  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY136372; AAM97038.1; --  
DR GO; GO:0008757; P:S-adenosylmethionine-dependent methyltransferase; IEA.  
DR GO; GO:0016740; P:transferase activity; IEA.  
DR InterPro; IPR001601; Methyltransf.  
DR InterPro; IPR000051; SAM bind.  
KW Transferase; Methyltransferase.  
SQ SEQUENCE 376 AA; 42811 MW; 2B4F6BE2B112B3ED CRC64;

Query Match 5.5%; Score 26; DB 10; Length 376;  
Best Local Similarity 100.0%; Pred. No. 4.2e-18;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 259 MLDLKPQKQVLDVCGIGGDFYMAE 284  
Db 290 MLDLKPQKQVLDVCGIGGDFYMAE 315

RESULT 4  
Q84SA4 PRELIMINARY; PRT; 493 AA.  
AC Q84SA4; 24, Created)  
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE Phosphoethanolamine N-methyltransferase.  
OS Aster tripolium (Sea aster).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC Campanulids; Asterales; Asteraceae; Asteroideae; Asteraceae; Aster.  
OX NCBI\_TaxID=74787;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Takeda M., Uno Y., Kanachi M., Inagaki N.;  
RT "Analyze of nine cDNAs for salt-inducible genes in the halophyte sea aster."  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB090883; BACS7960.1; --  
DR GO; GO:0008757; P:S-adenosylmethionine-dependent methyltransferase; IEA.  
DR GO; GO:0016740; P:transferase activity; IEA.  
DR InterPro; IPR001601; Methyltransf.  
DR InterPro; IPR000051; SAM bind.  
DR Methyltransferase, Transferase.  
SQ SEQUENCE 493 AA; 56002 MW; 9P2C7369192B6DA5 CRC64;

Query Match 5.5%; Score 26; DB 10; Length 493;  
Best Local Similarity 100.0%; Pred. No. 5.4e-18;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 YPDNSFDVYSRDTILHIQDKPALFR 351  
Db 346 YPDNSFDVYSRDTILHIQDKPALFR 371

RESULT 5  
Q7XJ2

ID Q7XJ2 PRELIMINARY; PRT; 491 AA.  
 AC Q7XJ2; 01-OCT-2003 (TRENBLrel. 25, Created)  
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
 DE Phosphoethanolamine N-methyltransferase.  
 OS Brassica napus (rape).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Brassica.  
 OX NCBI\_TaxID=3708;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Ye C., Li J., Yang J., Wang B.;  
 RT "Plant gene cloning and expression analysis under different  
 RL stresses";  
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
 KW EMBL; AY319479; AAP3582.1; -;  
 FW Methyltransferase; Transferase.  
 SQ SEQUENCE 491 AA; 55938 MW; A08C3318737031EE CRC64;  
 Query Match 5.1%; Score 24; DB 10; Length 491;  
 Best Local Similarity 100.0%; Pred. No. 6.7e-16; Mismatches 0; Indels 0; Gaps 0;  
 Matches 24; Conservative 0;  
 QY 261 DLKPGQKVLVGGCGGDFYMAE 284  
 DB 279 DLKPGQKVLVGGCGGDFYMAE 302  
 RESULT 6  
 Q8VX1 PRELIMINARY; PRT; 498 AA.  
 AC Q8VX1  
 DT 01-MAR-2002 (TRENBLrel. 20, Created)  
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
 DE Phosphoethanolamine methyltransferase.  
 OS Triticum aestivum (wheat).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Triticeae; Triticum.  
 OX NCBI\_TaxID=4565;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Frenette Charron J.-B., Breton G., Danyluk J., Muzac I., Ibrahim R.,  
 RA Sarhan F.;  
 RT "Molecular and biochemical characterization of a cold regulated  
 RT phosphoethanolamine methyltransferase from wheat";  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY065971; AAL40895.1; -;  
 DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransferase.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR InterPro; IPR001601; Methyltransf.  
 DR InterPro; IPR000051; SAM bind.  
 KW Transferase; Methyltransferase.  
 SQ SEQUENCE 498 AA; 56858 MW; 37BB7134E2DEA148 CRC64;  
 Query Match 4.4%; Score 21; DB 10; Length 498;  
 Best Local Similarity 100.0%; Pred. No. 9.5e-13; Mismatches 0; Indels 0; Gaps 0;  
 Matches 21; Conservative 0;  
 QY 265 GQKVLVGGCGGDFYMAE 285  
 DB 291 GQKVLVGGCGGDFYMAE 311  
 RESULT 7  
 Q8LJ10 PRELIMINARY; PRT; 499 AA.  
 ID Q8LJ10  
 AC Q8LJ10  
 DT 01-OCT-2002 (TRENBLrel. 22, Created)  
 DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)

DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
 DE Putative phosphoethanolamine methyltransferase.  
 GN P0431H09.20.  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC  
 RT clone: P0431H09.20";  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF003248; BAC10708.1; -;  
 DR Gramineae; Oryzoideae; Oryza.  
 DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransferase.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR InterPro; IPR001601; Methyltransf.  
 DR InterPro; IPR000051; SAM bind.  
 KW Transferase; Methyltransferase.  
 SQ SEQUENCE 499 AA; 56786 MW; ADDF3D04E62D188E CRC64;  
 Query Match 3.6%; Score 17; DB 10; Length 499;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-08; Mismatches 0; Indels 0; Gaps 0;  
 Matches 17; Conservative 0;  
 QY 331 FDVYSRDTLHIOQKP 347  
 DB 358 FDVYSRDTLHIOQKP 374  
 RESULT 8  
 Q82720 PRELIMINARY; PRT; 346 AA.  
 ID Q82720  
 AC Q82720  
 DT 01-NOV-1998 (TRENBLrel. 08, Created)  
 DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
 DE S-adenosyl-methionine-sterol-C-methyltransferase.  
 OS Nicotiana tabacum (Common tobacco).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC Lamiales; Solanales; Solanaceae; Nicotiana.  
 OX NCBI\_TaxID=4097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=xanthi SH6;  
 RX MEDLINE=98417432; PubMed=9746350;  
 RA Bouvier-Nave P., Hueselstein T., Benveniste P.;  
 RT "Two families of sterol methyltransferases are involved in the first  
 RT and the second methylation steps of plant sterol biosynthesis";  
 RL Eur. J. Biochem. 256:189-96 (1998).  
 DR EMBL; U01312; AAC34951.1; -;  
 DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransferase.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR InterPro; IPR000051; SAM bind.  
 KW Methyltransferase; Transferase.  
 SQ SEQUENCE 346 AA; 38871 MW; E7A3F9534A9D94A4 CRC64;  
 Query Match 3.4%; Score 16; DB 10; Length 346;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-07; Mismatches 0; Indels 0; Gaps 0;  
 Matches 16; Conservative 0;  
 QY 262 LKPGQKVLVGGCGIGG 277  
 DB 98 LKPGQKVLVGGCGIGG 113  
 RESULT 9  
 Q82434 PRELIMINARY; PRT; 349 AA.  
 ID Q82434

O82434;  
AC 01-NOV-1998 (TRENBLrel. 08, Created)  
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)  
DE 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
DS S-adenosyl-methionine cycloartenol-C24-methyltransferase.  
OS Nicotiana tabacum (Common tobacco).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC Lamiales; Solanales; Solanaceae; Nicotiana.  
OX NCBI\_TaxID=4097;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=xanthi SH6; TISSUE=Callus;  
RX MEDLINE=98417432; PubMed=9746350;  
RA Bouvier-Nave P., Huselestein T., Benveniste P.;  
RT "Two families of sterol methyl-transferases are involved in the first  
and the second methylation steps of plant sterol biosynthesis.";  
RL Eur. J. Biochem. 256:88-96(1998).  
DR EMBL; AF053766; AAC35787.1; -;  
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransferase. . .; IEA.  
DR GO; GO:0016740; F:transferase activity; IEA.  
DR InterPro; IPR000051; SAM bind.  
KW Methyltransferase; Transferase.  
SQ SEQUENCE 349 AA; 39028 MW; ED6A5853CABDCD77 CRC64;  
  
Query Match 3.4%; Score 16; DB 10; Length 349;  
Best Local Similarity 100.0%; Pred. No. 1.2e-07;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 262 LKPGQKVLVDVGGGIGG 277  
DB 99 LKPGQKVLVDVGGGIGG 114  
  
RESULT 10  
OQ3445 PRELIMINARY; PRT; 367 AA.  
AC Q3445;  
DT 01-NOV-1996 (TRENBLrel. 01, Created)  
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
DE S-adenosyl-L-methionine:delta24-sterol-C-methyltransferase.  
OS Glycine max (Soybean).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
OX NCBI\_TaxID=3847;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Williams 82; TISSUE=Etisolated hypocotyl;  
RX MEDLINE=96199190; PubMed=8621604;  
RA Shi J., Gonzales R.A., Bhattacharya M.K.;  
RT "Identification and characterization of an S-adenosyl-L-methionine:  
delta 24-sterol-C-methyltransferase cDNA from soybean.";  
RL J. Biol. Chem. 271:9384-9389(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Williams 82; TISSUE=Etisolated hypocotyl;  
RA Clouse J.A.;  
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U43683; AAB04057.1; -;  
DR PIR; T06780; T06780.  
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransferase. . .; IEA.  
DR GO; GO:0016740; F:transferase activity; IEA.  
DR InterPro; IPR000051; SAM bind.  
KW Methyltransferase; Transferase.  
SQ SEQUENCE 367 AA; 41518 MW; B3BEP8D2F0B22FB4 CRC64;  
  
Query Match 3.4%; Score 16; DB 10; Length 367;  
Best Local Similarity 100.0%; Pred. No. 1.2e-07;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 262 LKPGQKVLVDVGGGIGG 277

Db 119 LKPGQKVLVDVGGGIGG 134  
  
RESULT 11  
O8LKM1 PRELIMINARY; PRT; 336 AA.  
AC O8LKM1;  
DT 01-OCT-2002 (TRENBLrel. 22, Created)  
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)  
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
DE Cephalopod.  
GN CPH.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX PubMed=12100483;  
RA Schrick K., Mayer U., Martin G., Bellini C., Kuhn C., Schmidt J.,  
RA Jurgens G.;  
RT "Interactions between sterol biosynthesis genes in embryonic  
development of Arabidopsis.";  
RL Plant J. 31:61-73(2002).  
DR EMBL; AF494289; AAMS3553.1; -;  
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransferase. . .; IEA.  
DR InterPro; IPR000051; SAM bind.  
SQ SEQUENCE 336 AA; 38267 MW; BF5324A1B439B174 CRC64;  
  
Query Match 3.0%; Score 14; DB 10; Length 336;  
Best Local Similarity 100.0%; Pred. No. 1.4e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 264 LKPGQKVLVDVGGGIGG 277  
DB 94 LKPGQKVLVDVGGGIGG 107  
  
RESULT 12  
O9LM02 PRELIMINARY; PRT; 336 AA.  
AC O9LM02;  
DT 01-OCT-2000 (TRENBLrel. 15, Created)  
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
DE SAM:cycloartenol-C24-methyltransferase (24-sterol C-methyltransferase)  
(Sterol methyltransferase SMT1).  
GN SMT1 OR ATSG13710.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RA Schaeffer A., Schaller H., Benveniste P.;  
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Columbia;  
RX MEDLINE=98069011; PubMed=9405937;  
RA Kotani H., Nakamura Y., Sato S., Kaneko T., Asamizu E., Miyajima N.,  
RA Tabata S.;  
RT "Structural analysis of Arabidopsis thaliana chromosome 5. II.  
Sequence features of the regions of 1,044,062 bp covered by thirteen  
DNA res. 4:291-300(1997)."  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;

RA Diener A.C., Li H., Zhou W.-X., Whoriskey W.J., Nes W.D., Fink G.R.;  
 RT "Effects of reduced C-24 sterol alkylations on plant growth due to a  
 RL deficiency in sterol methyltransferase 1";  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 (4)  
 RP SEQUENCE FROM N.A.  
 RA Southwick A., Karlin-Neumann G., Nguyen M., Tripp M., Miranda M.,  
 RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,  
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,  
 RA Kim C., Lin J., Liu S.X., Naruaka M., Pham P.K., Sakano H.,  
 RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,  
 RA Ecker J., Theologis A., Davis R.W.;  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 (5)  
 RP SEQUENCE FROM N.A.  
 RA Nguyen M., Karlin-Neumann G., Southwick A., Tripp M., Miranda M.,  
 RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,  
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,  
 RA Kim C., Lin J., Liu S.X., Naruaka M., Pham P.K., Sakano H.,  
 RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,  
 RA Ecker J., Theologis A., Davis R.W.;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF090372; AAC78847.1; -  
 DR EMBL; AB006704; BAB08698.1; -  
 DR EMBL; AF195648; AAC28462.1; -  
 DR EMBL; AY120716; AAM53374.1; -  
 DR EMBL; BT000058; AAN15377.1; -  
 DR GO; GO:0008757; P:S-adenosylmethionine-dependent methyltransf. . . ; IEA.  
 DR GO; GO:0016740; P:transferase activity; IEA.  
 DR InterPro; IPR000051; SAM bind.  
 KW Methyltransferase; Transferase.  
 SQ SEQUENCE 336 AA; 3268 MW; 4649B3B68DEICE9 CRC64;  
 Query Match 3.0%; Score 14; DB 10; Length 336;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 264 PQQKVLVDGCGIGG 277  
 DB 94 PQQKVLVDGCGIGG 107  
 RESULT 13  
 O24328 PRELIMINARY; PRT; 346 AA.  
 ID O24328;  
 AC O24328;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE 8-adenosyl-methionine-sterol-C-methyltransferase.  
 OS Ricinus communis (Castor bean)  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosoids I; Malpighiales; Euphorbiaceae; Acalyphaceae;  
 OC Ricinus.  
 OX NCBI\_TaxID=3988;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Kaker 296;  
 RX MEDLINE=97352551; PubMed=9708946;  
 RA Bouvier-Nave P., Huselestein T., Despres T., Benveniste P.;  
 RT "Identification of cDNAs encoding sterol methyltransferases involved  
 RL in the second methylation step of plant sterol biosynthesis";  
 RL Eur. J. Biochem. 246:518-529 (1997).  
 DR EMBL; U01313; AAB62812.1; -  
 DR EMBL; T10173; T10173.  
 DR GO; GO:0008757; P:S-adenosylmethionine-dependent methyltransf. . . ; IEA.  
 DR GO; GO:0016740; P:transferase activity; IEA.  
 DR InterPro; IPR000051; SAM bind.  
 KW Methyltransferase; Transferase.  
 SQ SEQUENCE 346 AA; 38930 MW; 96F3DA79CA81AE82 CRC64;  
 Query Match 2.5%; Score 12; DB 10; Length 346;

Best Local Similarity 100.0%; Pred. No. 0.0018;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 266 QKVLVDGCGIGG 277  
 DB 102 QKVLVDGCGIGG 113  
 RESULT 14  
 Q84M50 PRELIMINARY; PRT; 330 AA.  
 ID Q84M50;  
 AC Q84M50;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Putative endosperm C-24 sterol methyltransferase.  
 GN OSJNBA0059E14.10.  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ccv. Nipponbare;  
 RA Buell C.R., Yuan Q., Cuyang S., Liu J., Gansberger K., Jones K.M.,  
 RA Overton II L.L., Taitzin T., Kim M.M., Bera J.J., Jin S.S.,  
 RA Fadroah D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,  
 RA Vanaken S.S., Riedmuller S.B., Utterback T.F., Feldblyum T.V.,  
 RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,  
 RA White O., Salzberg S.L., Fraser C.M.;  
 RT "Oryza sativa chromosome 3 BAC OSJNBA0059E14 genomic sequence";  
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
 (2)  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ccv. Nipponbare;  
 RA Buell R.;  
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC135958; AAP21419.1; -  
 DR GO; GO:0008757; P:S-adenosylmethionine-dependent methyltransf. . . ; IEA.  
 DR GO; GO:0016740; P:transferase activity; IEA.  
 DR InterPro; IPR000051; SAM bind.  
 KW Methyltransferase; Transferase.  
 SQ SEQUENCE 330 AA; 36900 MW; 1ADC1039C3E73C7D CRC64;  
 Query Match 2.3%; Score 11; DB 10; Length 330;  
 Best Local Similarity 100.0%; Pred. No. 0.02;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 267 KVLVDGCGIGG 277  
 DB 85 KVLVDGCGIGG 95  
 RESULT 15  
 P93852 PRELIMINARY; PRT; 344 AA.  
 ID P93852;  
 AC P93852;  
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Endosperm C-24 sterol methyltransferase.  
 GN ESMT1  
 OS Zea mays (Maize)  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC PACCAD clade; Panicoideae; Andropogoneae; Zea.  
 OX NCBI\_TaxID=4577;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97435974; PubMed=9290641;  
 RA Grebenok R.J., Galbraith D.W., Penna D.D.;  
 RT "Characterization of Zea mays endosperm C-24 sterol methyltransferase;



RT one of two types of sterol methyltransferase in higher plants.";  
RL Plant Mol. Biol. 34:891-896(1997).  
DR EMBL; U79669; AAB70886.1; -.  
DR PIR; T04138; T04138.  
DR GO; GO:0008757; P:S-adenosylmethionine-dependent methyltransf. . .; IEA.  
DR GO; GO:0016740; P:transferase activity; IEA.  
DR InterPro; IPR000051; SAM bind.  
KW Methyltransferase; Transferase.  
SQ SEQUENCE 344 AA; 38779 MW; 2794A45858E29EAE CRC64;  
  
Query Match 2.3%; Score 11; DB 10; Length 344;  
Best Local Similarity 100.0%; Pred. NO. 0.021;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 267 KVLGVGGGIGG 277  
|||||  
DB 103 KVLGVGGGIGG 113

Search completed: July 26, 2004, 13:23:30  
Job time : 43 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2004, 13:10:26 ; Search time 54 seconds  
(without alignments)  
2474.907 Million cell updates/sec

Title: US-10-031-331B-40  
Perfect score: 2515  
Sequence: 1 HTVDLTIHMLDLSQASDLID.....KLRRTAGEQWGLFVAKK 473

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: A\_Geneseq\_29Jan04:\*

- 1: Geneseq1980e:\*
- 2: Geneseq1990a:\*
- 3: Geneseq2000b:\*
- 4: Geneseq2001s:\*
- 5: Geneseq2002s:\*
- 6: Geneseq2003as:\*
- 7: Geneseq2003bs:\*
- 8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2515	100.0	473	4	AAB80627
2	2204	87.6	494	4	AAB80627
3	1829.5	72.7	491	4	AAB80627
4	1270	50.5	289	4	AAB80627
5	1033	41.1	287	7	ABM74462
6	779	31.0	191	7	ABM74064
7	658	26.2	168	4	AAB80627
8	241.5	9.6	281	6	AAB35493
9	225.5	9.0	276	4	ABG99864
10	224	8.9	280	6	ABP96395
11	224	8.9	280	7	ADD19156
12	218.5	8.7	283	3	AAB32507
13	218.5	8.7	283	7	ADD10279
14	214	8.5	280	6	ABP96396
15	214	8.5	280	7	ADD19157
16	212.5	8.4	346	5	AB13618
17	210.5	8.4	317	2	AAM95016
18	210.5	8.4	317	6	ABP96397
19	209.5	8.3	285	3	AAY70038
20	200.5	8.0	279	3	AAY70041
21	200.5	8.0	345	6	ABP96378
22	200.5	8.0	345	7	ADD19145
23	199.5	7.9	283	6	ABP57689
24	198	7.9	344	3	AB160392
25	194.5	7.7	283	2	AAY39308

26	194.5	7.7	299	4	AAB70954
27	191	7.6	299	6	ABP96392
28	190	7.6	299	6	ABM66984
29	189.5	7.5	376	6	ABP96379
30	189.5	7.5	376	7	ADD19146
31	187.5	7.5	330	6	ABP96399
32	187.5	7.5	336	3	AAG21651
33	187.5	7.5	339	3	AAG21525
34	187.5	7.5	361	3	AAG21524
35	187.5	7.5	366	3	AAG21650
36	187.5	7.5	370	4	AAB99888
37	187	7.4	293	3	AAG21652
38	187	7.4	296	3	AAG21526
39	186.5	7.4	314	5	ABP93454
40	186.5	7.4	315	6	ABP96391
41	186.5	7.4	336	3	AAG45951
42	186.5	7.4	336	3	AAG45954
43	186.5	7.4	358	3	AAG45950
44	186.5	7.4	365	3	AAG45953
45	186	7.4	293	3	AAG45955

## ALIGNMENTS

RESULT 1  
AAB80627  
ID AAB80627 standard; protein; 473 AA.

XX AAB80627;  
AC  
XX  
DT 06-AUG-2003 (revised)  
DT 02-MAY-2001 (first entry)  
XX  
DE Environmental stress tolerant protein SEQ ID 40.  
XX  
XX Environmental stress resistance; salt; heat; desert; transgenic plant.  
OS Suaeda japonica.  
XX  
XX WO200106006-A1.  
XX  
PD 25-JAN-2001.  
XX  
PF 19-JUL-2000; 2000WO-JP004862.  
XX  
PR 19-JUL-1999; 99JP-00235910.  
PR 24-MAR-2000; 2000JP-00085377.  
XX  
(NISC-) JAPAN SCI & TECHNOLOGY CORP.  
Yamada A, Ozeki Y, Saito T;  
WPI: 2001-147355/15.  
N-PSDB; AAF74206.

Screening method to obtain DNA encoding environmental stress resistance factor, useful for producing transgenic plants resistant to environmental stress.

Claim 64; Page 125-127; 167pp; Japanese.

Polynucleotide sequences AAF74187 - AAF74218 encode proteins AAB80608 - AAB80639, which impart environmental stress resistance. The invention relates to a method for identifying DNA encoding proteins imparting environmental stress resistance. The method comprises inserting cDNA from a library originating in a salt-resistant organism into a host cell, culturing the transformants under conditions in which the untransformed host does not grow well, and selecting for viable clones. The method is useful for obtaining DNA encoding environmental stress resistance factors. The DNA encoding proteins conferring environmental stress resistance can be used in the production of plants resistant to environmental stress, which can be cultivated in unfavourable

CC environments such as deserts, salt damaged ground, cold regions and the  
 CC oceans. They can be used for increasing the area of land covered by green  
 CC plants, and desert greening and afforestation, in order to counter the  
 CC effects of the increase in atmospheric carbon dioxide concentration. PCR  
 CC primers AAF74219 and AAF74220 are used in an example illustrating the  
 CC method of the invention. (Updated on 06-AUG-2003 to correct OS field.)  
 XX  
 SQ Sequence 473 AA;

Query Match 100.0%; Score 2515; DB 4; Length 473;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-237; Indels 0; Gaps 0;  
 Matches 473; Conservative 0; Mismatches 0;  
 QY 1 HTVDLTIAEMMLDSQASDLKERRPEITLSMLPPELGKCLLELGAGIGRTGLAEKAGQV 60  
 DB 1 HTVDLTIAEMMLDSQASDLKERRPEITLSMLPPELGKCLLELGAGIGRTGLAEKAGQV 60  
 QY 61 IALDFTESAIKKNVEINGHYKNVKNPCADVTSTLSPPHSLDVIPIFNWMLMYLSDERVE 120  
 DB 61 IALDFTESAIKKNVEINGHYKNVKNPCADVTSTLSPPHSLDVIPIFNWMLMYLSDERVE 120  
 QY 121 NLVERMLKWLKPGGYIPFRESCTHQSGDHKRNKNTHTYRPPRYTKAFKCHLQDGSNS 180  
 DB 121 NLVERMLKWLKPGGYIPFRESCTHQSGDHKRNKNTHTYRPPRYTKAFKCHLQDGSNS 180  
 QY 181 YELSLSCCKTCGAVVRNKNQNOISMLWQVDSKDKGFORFLDTSTQYKCNILRYERVF 240  
 DB 181 YELSLSCCKTCGAVVRNKNQNOISMLWQVDSKDKGFORFLDTSTQYKCNILRYERVF 240  
 QY 241 GPGYVSTGGYETTKFVSMILDKPGQKVLVGGGIGGDFYMAETFDVWVGFDSLVSNNI 300  
 DB 241 GPGYVSTGGYETTKFVSMILDKPGQKVLVGGGIGGDFYMAETFDVWVGFDSLVSNNI 300  
 QY 301 SPALERSIGLKCAVEFEVADCTKINYPDNSFDVIYSRDTTLIHIOQKPAFLFRSYKWLKPG 360  
 DB 301 SPALERSIGLKCAVEFEVADCTKINYPDNSFDVIYSRDTTLIHIOQKPAFLFRSYKWLKPG 360  
 QY 361 GKVLSDYCKKAGPPSPEFAAYIKORGVDLHVKEVGMKQKAGFDVVLAEORTQPIRV 420  
 DB 361 GKVLSDYCKKAGPPSPEFAAYIKORGVDLHVKEVGMKQKAGFDVVLAEORTQPIRV 420  
 QY 421 LRKELETVEKEDVIFSDSEEDYNDIVGGNDKLRRTAKGEQRMGLFVAKKX 473  
 DB 421 LRKELETVEKEDVIFSDSEEDYNDIVGGNDKLRRTAKGEQRMGLFVAKKX 473

## RESULT 2

AAE09760  
 ID AAE09760 standard; protein; 494 AA.

XX  
 AC AAE09760;

XX 29-NOV-2001 (first entry)

XX Spinach PEAMT protein.

XX Spinach; S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase;  
 KW PEAMT; cellular intermediate; phospho-dimethyl ethanolamine; choline;  
 KW phosphomono-methyl ethanolamine; phosphocholine; phosphatidylcholine;  
 KW glycine betaine; choline-O-sulphate; lipid content alteration;  
 KW osmotic stress tolerance; nutritional value; transgenic plant;  
 KW cryoprotectant.

XX Spinacia oleracea.

XX Key Location/Qualifiers

XX Key

XX Misc-difference 462

XX Note= "This residue is given as Lys in the sequence  
 shown as SEQ ID NO: 2 in figure 3 of the specification"

XX W0200168870-A2.

XX 20-SEP-2001.

XX 15-MAR-2001; 2001WO-US008352.  
 XX  
 XX 15-MAR-2000; 2000US-00525885.  
 XX (UYFL) UNIV FLORIDA.  
 XX (UYCA-) UNIV CARNEGIE MELLON.  
 XX  
 XX Hanson AD, Nuccio ML, Henry SA;  
 XX WPI; 2001-565796/63.  
 XX N-PSDB; AAD16797.  
 XX  
 XX New plant S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase  
 XX polypeptides, useful for modulating the levels of cellular intermediates  
 XX such as phosphodimethyl ethanolamine and for altering the lipid content in  
 XX plants cells.

XX Claim 1; Page 109; 150pp; English.

XX The present sequence is S-adenosyl-L-methionine:phosphoethanolamine N-  
 XX methyltransferase (PEAMT) protein from spinach. The PEAMT sequences are  
 XX useful for modulating the levels of cellular intermediates such as  
 XX phosphodimethyl ethanolamine, phosphomono-methyl ethanolamine, choline,  
 XX phosphocholine, phosphatidylcholine, choline-O-sulphate or glycine  
 XX betaine. They are useful for altering the lipid content in plant cells.  
 XX The polynucleotides are also useful for improving the osmotic stress  
 XX tolerance of a plant and increasing the cryoprotectant properties of a  
 XX plant. The present invention also relates to methods and compositions  
 XX comprising PEAMT used for generating transgenic plants with increased  
 XX nutritional value

XX Sequence 494 AA;

Query Match 87.6%; Score 2204; DB 4; Length 494;  
 Best Local Similarity 86.4%; Pred. No. 1.3e-206;  
 Matches 408; Conservative 34; Mismatches 30; Indels 0; Gaps 0;

QY 1 HTVDLTIAEMMLDSQASDLKERRPEITLSMLPPELGKCLLELGAGIGRTGLAEKAGQV 60  
 DB 22 HSVDLTIAEMMLDSQASDLKERRPEITLSMLPPELGKCLLELGAGIGRTGLAEKAGQV 81  
 QY 61 IALDFTESAIKKNVEINGHYKNVKNPCADVTSTLSPPHSLDVIPIFNWMLMYLSDERVE 120  
 DB 82 IALDFTESAIKKNVEINGHYKNVKNPCADVTSTLSPPHSLDVIPIFNWMLMYLSDERVE 141  
 QY 121 NLVERMLKWLKPGGYIPFRESCTHQSGDHKRNKNTHTYRPPRYTKAFKCHLQDGSNS 180  
 DB 142 NLVERMLKWLKPGGYIPFRESCTHQSGDHKRNKNTHTYRPPRYTKAFKCHLQDGSNS 201  
 QY 181 YELSLSCCKTCGAVVRNKNQNOISMLWQVDSKDKGFORFLDTSTQYKCNILRYERVF 240  
 DB 202 YELSLSCCKTCGAVVRNKNQNOISMLWQVDSKDKGFORFLDTSTQYKCNILRYERVF 261  
 QY 241 GPGYVSTGGYETTKFVSMILDKPGQKVLVGGGIGGDFYMAETFDVWVGFDSLVSNNI 300  
 DB 262 GPGYVSTGGYETTKFVSMILDKPGQKVLVGGGIGGDFYMAETFDVWVGFDSLVSNNI 321  
 QY 301 SPALERSIGLKCAVEFEVADCTKINYPDNSFDVIYSRDTTLIHIOQKPAFLFRSYKWLKPG 360  
 DB 322 SPALERSIGLKCAVEFEVADCTKINYPDNSFDVIYSRDTTLIHIOQKPAFLFRSYKWLKPG 381  
 QY 361 GKVLSDYCKKAGPPSPEFAAYIKORGVDLHVKEVGMKQKAGFDVVLAEORTQPIRV 420  
 DB 382 GKVLSDYCKKAGPPSPEFAAYIKORGVDLHVKEVGMKQKAGFDVVLAEORTQPIRV 441  
 QY 421 LRKELETVEKEDVIFSDSEEDYNDIVGGNDKLRRTAKGEQRMGLFVAKKX 472  
 DB 442 LRKELETVEKEDVIFSDSEEDYNDIVGGNDKLRRTAKGEQRMGLFVAKKX 493

RESULT 3  
 AAB99889



XX SQ Sequence 289 AA;  
 Query Match 50.5%; Score 1270; DB 4; Length 289;  
 Best Local Similarity 88.7%; Pred. No. 2e-115;  
 Matches 235; Conservative 15; Mismatches 15; Indels 0; Gaps 0;

QY 1 HTVDLTIAEMMLDSQADLDKPERPILSMPLPCKKLELGGAGIGRTGELAKAGQV 60  
 DB 22 HSDVLTIAEMMLDSQADLDKPERPILSMPLPCKKLELGGAGIGRTGELAKAGQV 81  
 QY 61 IALDFIRSAIKKNEVINGHYKVKFMCADVTSPFLSPHSLDVIIFSNWLLMYLSDEVE 120  
 DB 82 IALDFIRSAIKKNEVINGHYKVKFMCADVTSPFLSPHSLDVIIFSNWLLMYLSDEVE 141  
 QY 121 NLVERMLKWLKPGGYIFFRSCFHQSGDHKKNPHTHYREPRFYTKAFKCHLODGSNS 180  
 DB 142 NLVERMLKWLKPGGYIFFRSCFHQSGDHKKNPHTHYREPRFYTKAFKCHLODGSNS 201  
 QY 181 YELSLSCCKCIGAVKKNQNOISWLWQVDSKDDKGFORFLDTSQYKNSILRYRVP 240  
 DB 202 YELSLSCCKCIGAVKKNQNOISWLWQVDSKDDKGFORFLDTSQYKNSILRYRVP 261  
 QY 241 GFGYVSTGGYETTKFVSMLDLKP 265  
 DB 262 GFGYVSTGGYETTKFVSMLDLKP 286

RESULT 5  
 ABM74462  
 ID ABM74462 standard; protein; 287 AA.  
 AC ABM74462;  
 XX  
 DT 17-OCT-2003 (first entry)  
 DE DNA clone originating in barley containing SNP sequence #872.  
 KW Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis.  
 OS Hordeum vulgare.  
 XX  
 PN WO2003057877-A1.  
 PD 17-JUL-2003.  
 XX  
 PF 16-DEC-2002; 2002WO-IB005403.  
 XX  
 PR 20-DEC-2001; 2001JP-00387059.  
 PR 20-DEC-2001; 2001JP-00387131.  
 PR 20-DEC-2001; 2001JP-00403299.  
 PR 20-DEC-2001; 2001JP-00403300.  
 PR 27-SEP-2002; 2002JP-00327515.  
 XX  
 PA (UTNI-) UNIV JAPAN OKAYAMA.  
 XX  
 PI Sato K, Takeda K, Kohara Y;  
 XX  
 DR WPI; 2003-587127/55.  
 XX  
 PT Single nucleotide polymorphism sites in barley varieties and DNA  
 PT sequences containing them for analysis and identification of barley  
 PT varieties and production of barley transformants with desired  
 PT characteristics.  
 XX  
 PS Disclosure; SEQ ID XX; 284pp; Japanese.  
 XX  
 CC The present invention relates to oligonucleotide clones originating in  
 CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms  
 CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley  
 CC varieties, identification of particular varieties and genotype-phenotype  
 CC analysis, isolation of specific genes and creation of new varieties by  
 CC transformation of barley varieties with them and production of new barley

CC varieties with desired properties. The present sequence represents an  
 CC oligonucleotide clone sequence featured in the specification. The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published-pct-sequences  
 XX  
 SQ Sequence 287 AA;  
 Query Match 41.1%; Score 1033; DB 7; Length 287;  
 Best Local Similarity 76.9%; Pred. No. 3.3e-92;  
 Matches 193; Conservative 24; Mismatches 32; Indels 0; Gaps 0;

QY 1 HTVDLTIAEMMLDSQADLDKPERPILSMPLPCKKLELGGAGIGRTGELAKAGQV 60  
 DB 37 HSDVLTIAEMMLDSQADLDKPERPILSMPLPCKKLELGGAGIGRTGELAKAGQV 96  
 QY 61 IALDFIRSAIKKNEVINGHYKVKFMCADVTSPFLSPHSLDVIIFSNWLLMYLSDEVE 120  
 DB 97 IALDFIRSAIKKNEVINGHYKVKFMCADVTSPFLSPHSLDVIIFSNWLLMYLSDEVE 156  
 QY 121 NLVERMLKWLKPGGYIFFRSCFHQSGDHKKNPHTHYREPRFYTKAFKCHLODGSNS 180  
 DB 157 NLVERMLKWLKPGGYIFFRSCFHQSGDHKKNPHTHYREPRFYTKAFKCHLODGSNS 216  
 QY 181 YELSLSCCKCIGAVKKNQNOISWLWQVDSKDDKGFORFLDTSQYKNSILRYRVP 240  
 DB 217 YELSLSCCKCIGAVKKNQNOISWLWQVDSKDDKGFORFLDTSQYKNSILRYRVP 276  
 QY 241 GFGYVSTGGYETTKFVSMLDLKP 251  
 DB 277 GFGYVSTGGYETTKFVSMLDLKP 287

RESULT 6  
 ABM74064  
 ID ABM74064 standard; protein; 191 AA.  
 AC ABM74064;  
 XX  
 DT 17-OCT-2003 (first entry)  
 DE DNA clone originating in barley containing SNP sequence #474.  
 KW Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis.  
 OS Hordeum vulgare.  
 XX  
 PN WO2003057877-A1.  
 PD 17-JUL-2003.  
 XX  
 PF 16-DEC-2002; 2002WO-IB005403.  
 XX  
 PR 20-DEC-2001; 2001JP-00387059.  
 PR 20-DEC-2001; 2001JP-00387131.  
 PR 20-DEC-2001; 2001JP-00403299.  
 PR 20-DEC-2001; 2001JP-00403300.  
 PR 27-SEP-2002; 2002JP-00327515.  
 XX  
 PA (UTNI-) UNIV JAPAN OKAYAMA.  
 XX  
 PI Sato K, Takeda K, Kohara Y;  
 XX  
 DR WPI; 2003-587127/55.  
 XX  
 PT Single nucleotide polymorphism sites in barley varieties and DNA  
 PT sequences containing them for analysis and identification of barley  
 PT varieties and production of barley transformants with desired  
 PT characteristics.  
 XX  
 PS Disclosure; SEQ ID XX; 284pp; Japanese.  
 XX  
 CC The present invention relates to oligonucleotide clones originating in

CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms  
 CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley  
 CC varieties, identification of particular varieties and genotype-phenotype  
 CC analysis, isolation of specific genes and creation of new varieties by  
 CC transformation of barley varieties with them and production of new barley  
 CC varieties with desired properties. The present sequence represents an  
 CC oligonucleotide clone sequence featured in the specification. The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published-pct-sequences  
 XX  
 XX Sequence 191 AA;

Query Match 31.08; Score 779; DB 7; Length 191;  
 Best Local Similarity 73.84; Pred. NO. 1.4e-67;  
 Matches 141; Conservative 26; Mismatches 24; Indels 0; Gaps 0;  
 QY 282 MAETFDVEVGFDLSVNMISFALERSIGLKCAVEFEVADCTKINYPDINSFDVIYSRDTIL 341  
 DB 1 MAENYDVHVVGIDLSINNVSFALHAIGRKA VEFEVADCTTKTYPDNTFDVIYSRDTIL 60  
 QY 342 HIQDKPALFRSPYKWLPGGKVLISDYCKKAGPPSPFAAYIKQRYGDLHDVKEYGWLK 401  
 DB 61 HIQDKPALFRSPYKWLPGGKVLISDYCKSPGPEEFASYIKQRYGDLHDVKEYGWLK 120  
 QY 402 DAGFVDVLAERTQFIRVLKSLKETVEKEKDVISDFSEEDYNDIVGWNDKLRTAKG 461  
 DB 121 NAGFHDVTAERDSQFLKVLQRLAEVKEKDDFLADFGQEDYDIDVTGMNAKLQSSAG 180  
 QY 462 EQRWGLFVAKK 472  
 DB 181 EQRWGLFVGTK 191

RESULT 7  
 AAB99873  
 ID AAB99873 standard; protein; 168 AA.  
 AC AAB99873;  
 XX  
 DT 21-SEP-2001 (first entry)  
 XX  
 DE Physcomitrella patens 7a\_ppprotcl\_092\_e12rev protein.  
 XX  
 XX Tocopherol and carotenoid metabolism related protein; TCWRP; synthesis;  
 KM Physcomitrella patens; moss; algae; microorganism; fungus; plant;  
 KM identification; genome mapping; modulation; evolutionary study;  
 KM cellular production; fine chemical.  
 XX  
 OS Physcomitrella patens.  
 XX  
 PN W0200144276-A2.  
 XX  
 PD 21-JUN-2001.  
 XX  
 XX 14-DEC-2000; 2000WO-EP012698.  
 XX  
 XX 16-DEC-1999; 99US-0171121P.  
 XX  
 PA (BADI) BASF PLANT SCI GMBH.  
 XX  
 PI Lerchl J, Benz A, Ehrhardt T, Reindl A, Cirpus P, Bischoff P;  
 PI Frank M, Freund A, Duwenig E, Schmidt R, Reski R, Badur R;  
 XX  
 DR WPI; 2001-398121/42.  
 DR N-PSDB; AAB44246.  
 XX  
 PT Tocopherol and carotenoid metabolism related protein (TCWRP), used to  
 PT produce fine chemicals, is isolated from mosses, algae, microorganisms,  
 PT fungi, plants, or their fragments.  
 XX  
 XX Claim 28; Page 118-119; 123pp; English.

CC The present invention describes isolated tocopherol and carotenoid  
 CC metabolism related proteins (TCWRP) (i) from mosses or algae,  
 CC microorganisms or fungi, plants, or its fragments. (i) can be used as  
 CC enzymes in the production of fine chemicals or in the metabolism of  
 CC tocopherols and carotenoids. (i) also assist in transmembrane transport.  
 CC The fine chemicals that can be produced include lipids, fatty acids,  
 CC vitamins, cofactors, enzymes, amino acids, and nucleotide bases.  
 CC Nucleotide sequences, proteins, vectors and host cells from the present  
 CC invention can be used: (a) to identify mosses related to Physcomitrella  
 CC patens; (b) in mapping genomes of mosses related to Physcomitrella  
 CC ; (c) in the modulation of TCWRP activity; (d) in evolutionary studies;  
 CC (e) in the determination of functional TCWRP regions; (f) and in the  
 CC cellular production of fine chemicals. AAB44222 to AAB44262 encode the  
 CC Physcomitrella patens TCWRP proteins given in AAB99849 to AAB99889.  
 CC AAB44212 to AAB44221 represent nucleotide sequence used in the  
 CC exemplification of the present invention  
 XX  
 XX Sequence 168 AA;

Query Match 26.24; Score 650; DB 4; Length 168;  
 Best Local Similarity 72.94; Pred. NO. 8.1e-56;  
 Matches 121; Conservative 19; Mismatches 26; Indels 0; Gaps 0;  
 QY 307 SIGLKCAVEFEVADCTKINYPDINSFDVIYSRDTILHIQDKPALFRSPYKWLPGGKVLIS 366  
 DB 1 SIGLKCAVEFEVADCTKINYPDINSFDVIYSRDTILHIQDKPALFRSPYKWLPGGKVLIS 60  
 QY 367 DYCKKAGPPSPFAAYIKQRYGDLHDVKEYGWLKADGFDVLAERTQFIRVLKXLE 426  
 DB 61 DYCRAPQTPSAEFAAYIKQRYGDLHDVKEYGWLKADGFDVLAERTQFIRVLQRELA 120  
 QY 427 TVEKEKDVISDFSEEDYNDIVGWNDKLRTAKGQRWGLFVAKK 472  
 DB 121 TTEAGRODFINDFSEEDYNYIVSGMKSLKRCNSDEQKWLFIAYK 166

RESULT 8  
 AAB35493  
 ID AAB35493 standard; protein; 281 AA.  
 AC AAB35493;  
 XX  
 DT 17-JUN-2003 (first entry)  
 XX  
 DE Streptomyces platensis subspecies rosaceus dorrigocin ORF10 protein.  
 XX  
 KM Polyketide biosynthesis; dorrigocin; DORR; lactimidomycin; LACT.  
 XX  
 OS Streptomyces platensis.  
 XX  
 PN W0200288176-A2.  
 XX  
 PD 07-NOV-2002.  
 XX  
 XX 26-APR-2002; 2002WO-CA000591.  
 XX  
 XX 26-APR-2001; 2001US-0286346P.  
 XX  
 PA (ECOP-) ECOPIA BIOSCIENCES INC.  
 XX  
 PI Farnet CM, Zazopoulos E, Staffa A, Yang X;  
 XX  
 DR WPI; 2003-201222/19.  
 DR N-PSDB; AAD54217, AAD54227.  
 XX  
 PT Novel isolated or purified polypeptide involved in biosynthesis of  
 PT polyketide dorrigocin or polyketide lactimidomycin, useful for preparing  
 PT dorrigocin or lactimidomycin.  
 XX  
 XX Claim 13; Page 197-198; 312pp; English.  
 XX  
 CC The invention relates to novel proteins involved in the biosynthesis of  
 CC polyketide dorrigocin (DORR) or lactimidomycin (LACT) biosynthesis by

microorganisms. Sequences of the invention allow direct manipulation of dorrigoicin, lactimidomycin and related chemical structures via chemical engineering of the enzymes involved in the biosynthesis of dorrigoicin and lactimidomycin. They are useful for introducing chemical handles into normally inert positions that permit subsequent chemical modifications and facilitate the development of polyketides. The genes and proteins of the invention can also be used to generate a focused library of analogues around a polyketide lead candidate to fine-tune the compound for optimal properties. They are useful for generating antibodies specific for the polyketide biosynthesis. The present sequence is Streptomyces platensis subspecies rosaceus DORR ORF10 protein

XX Sequence 281 AA;

Query Match 9.6%; Score 241.5; DB 6; Length 281;  
Best Local Similarity 31.6%; Pred. No. 1.2e-14;  
Matches 65; Conservative 40; Mismatches 80; Indels 21; Gaps 7;  
QY 253 TKEFVSMDLKPGKVLVGGGCGGDFYMAETFDVYVQPLSVNMI--SPALERSIGL 310  
DB 63 TDTMTDLRIQGGQRVLVGGGCGGCPAMRIARTGAHVGTALSKDQIARATLABAGL 122  
QY 311 KCAVEFEVADCTKINYPNSFDVIYSRDTILHIQDKPALFRSYKWLKPGKVLISDYCK 370  
DB 123 SDREVEFHADAMELFPDDSDFAAIAIESIFHMPDRGRVLARIRVREGRLVLTDPFE 182  
QY 371 KAGPPSPFAAVIKQGYD--LHDV-----KEYGQMLKDG--FVDVLADRTQPIR 419  
DB 183 R-GFPAE-----KQPAVDRLLRDPINTLARPEYVPMLEADGLRVELL--DITEQSVR 234  
QY 420 VLKKELETVEKKEKQVFTSDSEEDYN 445  
DB 235 QTFFQMSQSGQEMQTVFDDAEKFS 260

RESULT 9  
ABG99864

ID ABG99864 standard; protein; 276 AA.

AC ABG99864;

XX 16-JAN-2003 (first entry)

DE S. cinnamomensis MonE/S-adenosylmethione-dependent methyltransferase.

XX Monensin; gene cluster; polyketide synthase; antibiotic; antihelminthic;  
XX insecticide; immunosuppressant; antifungal; antibacterial; polyether;  
XX mon B1; mon B11; mon C1; mon C11; mon H; mon R1; mon R11; mon T; mon AIX;  
XX mon AX.

XX Streptomyces cinnamomensis.

XX WO200168867-A1.

XX 20-SEP-2001.

XX 30-MAY-2000; 2000WO-GB002072.

XX 28-MAY-1999; 99GB-00012563.

XX (BIOT-) BIOTICA TECHNOLOGY LTD.

XX Leadlay PF, Staunton J, Oliynyk M;

XX WPI; 2001-611393/70.

XX N-PSDB; ABX04971.

XX New DNA sequence encoding polyketide synthase, useful for the production  
XX of polyketides such as antibiotic monensin.

XX Claim 6; Page 73; 21pp; English.

XX The invention relates to a DNA sequence which is a fully defined sequence

CC of 103551 base pairs appearing as ABX04971, or its variant, that it is  
CC not a sequence encoding all or part amino acids 1-920 encoded by mon A1  
CC as given in the specification. The DNA is the S. cinnamomensis polyketide  
CC antibiotic monensin biosynthetic gene cluster. Also included are a  
CC recombinant cloning or expression vector comprising the gene cluster, a  
CC transformed host cell which has been transformed to contain the gene,  
CC cluster (and is capable of expressing a corresponding polypeptide), a  
CC hybridization probe derived from the gene cluster (for identification and  
CC isolation of the same or analogous gene cluster, e.g. one which binds  
CC specifically to a region of the monensin gene cluster selected from mon  
CC B1, mon B11, mon C1, mon C11, mon H, mon R1, mon R11, mon T, mon AIX and  
CC mon AX), the use of the mon R1 gene or variant and a monensin promoter to  
CC control expression of a heterologous gene in Streptomyces cinnamomensis,  
CC a polypeptide encoded by a portion of the monensin gene cluster  
CC (preferably comprising mon B1, mon B11, mon AIX or mon AX or their  
CC mutants, alleles or variants), an epoxide enzyme encoded by mon C1, a  
CC cyclase enzyme encoded by mon C11, producing S. cinnamomensis capable of  
CC enhanced levels of production of monensin comprising engineering it to  
CC overexpress the mon R1 gene, S. cinnamomensis containing multiple copies  
CC of the mon R1 gene and/or its variants, expressing a gene heterologous to  
CC S. cinnamomensis comprising transforming S. cinnamomensis with DNA  
CC encoding a heterologous gene and expressing the gene under control of the  
CC activator gene mon R1 or actII/orf4 and 13-propyl erythromycin A. The  
CC processes and materials (enzyme systems, nucleic acids and vectors) are  
CC useful for preparing polyketides by recombinant synthesis. The  
CC polyketides are useful as insecticides, antibiotics, antihelminthics,  
CC antifungals, antibacterials or other pharmaceuticals. In particular the  
CC gene is useful for the production of monensin, an antibiotic polyether  
CC polyketide. The present sequence represents a protein encoded by the  
CC monensin gene cluster

XX Sequence 276 AA;

Query Match 9.0%; Score 225.5; DB 4; Length 276;

Best Local Similarity 28.5%; Pred. No. 4.3e-13;

Matches 63; Conservative 36; Mismatches 99; Indels 23; Gaps 5;

QY 243 GVSSTGG-----YETTKFVSMDLKPGKVLVGGGCGGDFYMAETFDVYVQ 292  
DB 35 GYVFDGEGQATFDENVMQMTDEMIRLDPADQDRVLDIGCGGTGPAQLARDEYVVG 94  
QY 293 FDLNVNMISFALERS--IGLKCAVEFEVADCTKINYPNSFDVIYSRDTILHIQDKPALF 350  
DB 95 TSVSARQVERGNRRBARAGLADRVRFQVDAMNLPDDGSGFDHCALESMLHMPDKQVVL 154  
QY 351 RSFYKWLKPGKVLISDYCKKAGPPS--PEFAAYIKQGY--DLHDVYKYGQMLKDGAFVDV 408  
DB 155 TEHRRVVKFGARMPFIADVMVYLNPDSPRPTATVSDTTIYAALTDIGDYDFIIRAAGMTVL 214  
QY 409 LAEDRTQPIRVLKKELETVEKKEKQVFTSDSEEDYNIVG 449  
DB 215 ELTDITRETATYDGYVEWIRAHND-----EYVDIIG 246

RESULT 10

ABP96395

ID ABP96395 standard; protein; 280 AA.

XX ABP96395;

XX 22-MAY-2003 (first entry)

XX Nostoc punctiforme gamma-tocopherol methyltransferase SEQ ID NO:39.

XX Gamma-tocopherol methyltransferase; methyltransferase; tocopherol;  
XX alpha-tocopherol; alpha-tocotrienol; stress; oxidative stress tolerance;  
XX oxygen; ozone; UV tolerance; ultra violet tolerance; cold tolerance;  
XX fungal pathogen; microbial pathogen; plant; enzyme.

XX Nostoc punctiforme.

XX WO2003016482-A2.





Db 5 LYQI-----QQFYDASSG-----LWQIWEHMHGYGADGTOKKDRQAQIDL 50  
 QY 253 TKEFVSMLDKPGQKVLNCGGCGDFYMAETFEVVGFDLSVNMISFALESII--GL 310  
 Db 51 IEELNAGVQNAEDILVCGGIGSSLYLAQFNKATGYLSPVQARATERALEANL 110  
 QY 311 KCAVEFEVADCTKINYPNSFDVYISDITLHIDQKALFRFYKWLKPGKVLISDYCK 370  
 Db 111 SLRTOFQVANAQAFADSDFLVMSLESGEHMPDKTFLQBCYRVLKPGKGLINVTWCH 170  
 QY 371 KAGPPSP 377  
 Db 171 RPTDESP 177

RESULT 12  
 AAB32507  
 ID AAB32507 standard; protein; 283 AA.  
 XX  
 AC AAB32507;  
 XX  
 XX 19-JAN-2001 (first entry)  
 XX  
 XX S. lavendulae Mit M encoded protein sequence.  
 XX  
 XX Mitomycin; biosynthesis; mitosome ring system; antibiotic; anti-cancer;  
 KW anti-inflammatory; immune-enhancer; immunosuppressant; asthma;  
 KW chronic obstructive pulmonary disease; respiratory inflammation;  
 KW fungicide; pesticide.  
 XX  
 XX Streptomyces lavendulae.  
 XX  
 XX WO200053737-A2.  
 XX  
 XX 14-SEP-2000.  
 XX  
 XX 10-MAR-2000; 2000WO-US006394.  
 XX  
 XX 12-MAR-1999; 99US-00266965.  
 XX  
 XX (MINU ) UNIV MINNESOTA.  
 PA (SHER/) SHERMAN D H.  
 PA (MAOY/) MAO Y.  
 PA (VARO/) VAROGLU M.  
 PA (HEMM/) HE M.  
 PA (SHEL/) SHELTON P C.  
 XX  
 XX Sherman DH, Mao Y, Varoglu M, He M, Sheldon PC;  
 XX  
 XX WPI; 2000-601980/57.  
 DR N-PSDB; AAC55802.  
 XX  
 XX Novel nucleic acid molecule comprising mitomycin biosynthetic gene  
 PT cluster useful for cloning mitomycin biosynthetic genes for elucidating  
 PT the molecular basis of mitosome ring system biosynthesis.  
 XX  
 XX Disclosure; Page 353; 399pp; English.  
 XX  
 XX This invention relates to isolated and purified nucleic acid molecules  
 CC from the mitomycin biosynthetic gene cluster. Mitomycins are a group of  
 CC natural products that contain a variety of functional groups, including  
 CC amino benzocoumarone and axiridine ring systems. The S. lavendulae  
 CC mitomycin biosynthetic gene cluster comprises 47 mitomycin genes spanning  
 CC 55kb of DNA. The invention includes an expression cassette comprising a  
 CC mitomycin biosynthetic gene operably linked to a promoter, and host cells  
 CC transformed with the cassette. The nucleotide, and protein sequences and  
 CC the transformed host cells of the invention result in antiasthmatic,  
 CC antiinflammatory, cytostatic, immunomodulatory, and antibiotic  
 CC activities. The nucleotide sequences are used to elucidate the molecular  
 CC basis for the biosynthesis of the mitosome ring system, as well as to  
 CC engineer the biosynthesis of novel natural products, e.g. antibiotics,  
 CC anti-inflammatory agents, anti-cancer agents, immune-enhancers,  
 CC immunosuppressants, agents to treat asthma, chronic obstructive pulmonary

CC disease as well as other disease involving respiratory inflammation, or  
 CC cholesterol-lowering agents or as crop protection agents (e.g. fungicides  
 CC or insecticides) as well as biopolymers, e.g., in packaging or biomedical  
 CC applications, or to engineer PHA monomer synthases. Sequences AAC55782-  
 CC C55881, AAC55815-C55849 and AAB32485-B32542 represent mitomycin  
 CC biosynthetic gene cluster DNA sequences and encoded proteins. Sequences  
 CC AAC55812-C55814, AAC55850-C55856 and AAC55862-C55869 represent PCR  
 CC primers used in the cloning of the mitomycin biosynthetic genes  
 XX  
 XX Sequence 283 AA;

Query Match 8.7%; Score 218.5; DB 3; Length 283;  
 Best Local Similarity 26.9%; Pred. No. 2.2e-12;  
 Matches 68; Conservative 46; Mismatches 92; Indels 47; Gaps 11;

QY 236 YERVFEGCVSTG-----GY-----ETTKFEVSM-----LDLKPGQKVLNCGGI 275  
 Db 23 YDRFTALGASLGENLHFGYWDSPDSQVFLAEYDDELTDMAERLRIGAGSRVLDLGGCV 82  
 QY 276 GGGDFYMAETFEVVGFDLSVNMISFALESIGLKCAVEFEVADCTKINYPNSFDV 333  
 Db 83 CTPGVRIARLSGAHVGTGISVSHQVVRANALAEAGLADRARFQRADAMDLPFDESDFA 142  
 QY 334 IYSRDTLHIDQKALFRFYKWLKPGKVLISDYCKAGPSPSPFAAYIKQGYDLHD- 392  
 Db 143 VIALESIIHMFDRQAQLVAGRVLRPGRLVLTDFPERA-PLAPEGRAAVQRY---LHDF 198  
 QY 393 -----VKEYQMLKDG-----FVDVLADRTQEFIVLRKELETVEKEKDVFIISDFSE 441  
 Db 199 MMTVMVSAEYPLLRGAGLWLEFLDI-SDQTEKTRLLS-----ERINSSKORLETQFG 254  
 QY 442 E-----DYNDIVG 449  
 Db 255 EMVNQFDPGLVG 267

## RESULT 13

ADE10279  
 ID ADE10279 standard; protein; 283 AA.

XX  
 AC ADE10279;

XX 29-JAN-2004 (first entry)

XX S. lavendulae mitomycin biosynthetic protein MitM.

XX Mitomycin biosynthetic protein; mitomycin C; antibiotic; MC; apoptosis;  
 KW tumour hypoxia; cytostatic; anti-tumour agent; cancer.

XX Streptomyces lavendulae.

XX US2003134398-A1.

XX 17-JUL-2003.

XX 12-SEP-2001; 2001US-00953348.

XX 12-SEP-2001; 2001US-00953348.

XX (SHER/) SHERMAN D H.

PA (MAOY/) MAO Y.

PA (VARO/) VAROGLU M.

PA (HEMM/) HE M.

PA (SHEL/) SHELTON P.

XX Sherman DH, Mao Y, Varoglu M, He M, Sheldon P;

XX WPI; 2003-863498/80.

DR N-PSDB; ADE10221.

XX New nucleic acid molecule comprising a sequence having mitomycin

PT biosynthetic gene cluster, useful for enhancing production of

PT antibiotics.

XX PS Disclosure; SEQ ID NO 109; 308pp; English.

XX CC The invention relates to an isolated and purified nucleic acid molecule comprising a sequence having mitomycin biosynthetic gene cluster, or its variant or fragment. Also included are an expression cassette comprising the nucleic acid molecule (operably linked to a promoter functional in a host cell), a recombinant bacterial host cell in which at least a portion of a nucleic acid molecule comprising mitomycin biosynthetic gene cluster is disrupted (resulting in a recombinant host cell that produces altered levels of mitomycin relative to a corresponding nonrecombinant bacterial host cell), introducing exogenous DNA into a refractory Streptomyces strain, identifying a nucleic acid molecule that is related to at least a portion of a nucleic acid molecule comprising a mitomycin gene cluster, preparing a compound or its salt from the recombinant host cell and a product produced by the recombinant host cell. The nucleic acid encodes a MitT, MitR, MitQ, MitP, MitO, MitN, MitM, MitL, MitK, MitJ, MitI, MitH, MitG, MitF, MitE, MitD, MitC, MitB, MitA and/or MnCA-MncY. The nucleic acid is useful for enhancing production of mitomycin antibiotics, which induce apoptosis and hence are useful as anti-tumour (via tumour hypoxia) agents and are useful in treating cancer. The gene cluster was isolated from Streptomyces lavendulae. The present sequence is a mitomycin biosynthetic protein of the invention.

XX SQ Sequence 283 AA;

Query Match 8.7%; Score 218.5; DB 7; Length 283;  
Best Local Similarity 26.9%; Pred. No. 2.2e-12;  
Matches 68; Conservative 46; Mismatches 92; Indels 47; Gaps 11;

QY 236 YERVPCGVSTG-----GY-----ETTKFVSM-----LDLKGQKVLGVGCI 275  
DB 23 YRFTALGAALGKLFHGFWDSPDSQVPLAETDLTWMARLIGNSRVLDLGGCV 82

QY 276 QGGDFYMAETDEVEVGGDLSVMI---SFALRSIGLKAVFEVADCTKYNPNNSDV 333  
DB 83 GTPGVRIARLGGKVTGSGVHEQVVRANALAEGLADRARFQRADMDLFFEDSFDA 142

QY 334 IYSRDTILHIQDKALFRSPYKMLKPGKVLISDYCKAGPPSPFAAYIKQGYDLHD- 392  
DB 143 VIALSIIHMPDRAQVLAQGVRLRPGRLVLTDFERR-PLAPEGRAVQRY---LHDP 198

QY 393 -----VKEYGOMLKDAG-----FVDVLAEDRTQFIRVLKELETVEKDVFSDFSE 441  
DB 199 MMTWSAENAYPPLKAGLWLEEFELDI-SQTLKTRFLLS---ERINSSKORLSTQGE 254

QY 442 E-----DYNDIVG 449  
DB 255 EMVNQPDGDLVG 267

RESULT 14  
ABP96396  
ID ABP96396 standard; protein; 280 AA.  
XX AC ABP96396;  
XX DT 22-MAY-2003 (first entry)  
XX DB Anabaena sp. gamma-tocopherol methyltransferase SEQ ID NO:40.  
XX KW Gamma-tocopherol methyltransferase; methyltransferase; tocopherol;  
XX KW alpha-tocopherol; alpha-tocotrienol; stress; oxidative stress tolerance;  
XX KW oxygen; ozone; UV tolerance; ultra violet tolerance; cold tolerance;  
XX KW fungal pathogen; microbial pathogen; plant; enzyme.  
XX OS Anabaena sp.  
XX FN WO2003016482-A2.  
XX PD 27-FEB-2003.  
XX PF 16-AUG-2002; 2002WO-US026047.

XX PR 17-AUG-2001; 2001US-0312758P.  
XX PA (MONS ) MONSANTO TECHNOLOGY LLC.  
XX PI Van Eenennaam A, Valentin HE, Karunanandaa B, Hao M, Aasen E;  
XX PI Levering C;  
XX WI 2003-269314/26.  
XX DR N-PSDB; ABZ80088.  
XX PT Novel substantially purified tocopherol protein and nucleic acid encoding  
XX PT the same, useful for producing a plant having a seed with increased alpha  
XX PT -tocopherol level.  
XX PS Claim 22; Page 197-198; 218pp; English.  
XX CC The present invention describes a substantially purified tocopherol  
XX CC protein (I). Also described is a method (M) for reducing expression of  
XX CC methyltransferase 1 (MT1) or gamma-tocopherol methyltransferase (GMT) in  
XX CC a plant, by transforming a plant with a nucleic acid molecule having an  
XX CC exogenous promoter region which functions in plant cells to cause the  
XX CC production of an mRNA molecule, and growing the transformed plant. (M) is  
XX CC useful for increasing the gamma-tocopherol content. A polynucleotide (II)  
XX CC encoding (I) can be used for producing a plant having a seed with an  
XX CC increased alpha-tocopherol or alpha-tocotrienol level. (II) is useful for  
XX CC modulating the reduction of the expression, expression, overexpression of  
XX CC (I) in a transformed plant to provide tolerance to a variety of stress,  
XX CC e.g. oxidative stress tolerance such as to oxygen or ozone, UV tolerance,  
XX CC cold tolerance, or fungal/microbial pathogen tolerance. (II) is useful  
XX CC for obtaining other nucleic acid molecules or homologues from the same  
XX CC species, to screen cDNA or genomic libraries, to isolated promoters of  
XX CC cell enhanced, cell specific, tissue enhanced, expression profiles, as  
XX CC developmentally or environmentally regulated expression profiles, as  
XX CC markers, for detecting single nucleotide polymorphisms, and to determine  
XX CC the level of (I) in a plant or pattern of expression of (I) encoded in  
XX CC part or whole by (II). The present sequence represents an Anabaena sp.  
XX CC gamma-tocopherol methyltransferase, from the present invention

XX SQ Sequence 280 AA;

Query Match 8.5%; Score 214; DB 6; Length 280;  
Best Local Similarity 28.9%; Pred. No. 5.9e-12;  
Matches 54; Conservative 35; Mismatches 68; Indels 30; Gaps 5;

QY 207 LKQKVDKDDKGFRLDTQYKCNILAYRVFGP-----GYVSTGGYETTKFVSMIDL 262  
DB 5 LYQOI-----QOYDASG-----LWEEIWMHMHGYGADGTQKRRQAQIDL 50

QY 263 -----XPGQKVLGVGCGIGCGDFYMAETDFVGVVFDLSVNMISFALERS--IGL 310  
DB 51 IEELLTWAGVQTAENILDVGGIGGSSLYLAGKLNKATGTLSPVQAARATERAKEAGL 110

QY 311 KCAVEFEVADCTKYNPNNSFDVIYSRDTILHIQDKPALFRSPYKMLKPGKVLISDYCK 370  
DB 111 SCRSQFLVANAQAMPFDNSFDLVNSLSEGEHMPDKTFLOECYVILKPGKLIWVTWCH 170

QY 371 KAGPESP 377  
DB 171 RPTDXTF 177

RESULT 15  
ADD19157  
ID ADD19157 standard; protein; 280 AA.  
XX AC ADD19157;  
XX DT 15-JAN-2004 (first entry)  
XX DE Anabaena sp. gamma-tocopherol methyl transferase protein.  
XX KW tocopherol biosynthesis pathway; methyltransferase; tocopherol;



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2004, 13:13:06 / Search time 19 Seconds  
(without alignments)  
1285.215 Million cell updates/sec

Title: US-10-031-331B-40

Perfect score: 2515

Sequence: 1 HTVDLTTEAMMLDSQASDLID.....KURRTAKEQRKGLFVAKKK 473

Scoring table: BLOSUM62

Gapop 10.0 / Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgm2\_6/ptodata/2/iaa/5A COMB.pdp:\*  
2: /cgm2\_6/ptodata/2/iaa/5B COMB.pdp:\*  
3: /cgm2\_6/ptodata/2/iaa/6A COMB.pdp:\*  
4: /cgm2\_6/ptodata/2/iaa/6B COMB.pdp:\*  
5: /cgm2\_6/ptodata/2/iaa/PCTRUS COMB.pdp:\*  
6: /cgm2\_6/ptodata/2/iaa/backfileal.pdp:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	218.5	8.7	283	4	US-09-266-965-109
2	217.5	8.6	363	3	US-09-041-718-5
3	210.5	8.4	317	4	US-09-118-637A-2
4	194.5	7.7	283	3	US-09-036-987A-13
5	194.5	7.7	283	3	US-09-370-700-13
6	194.5	7.7	283	4	US-09-603-207-13
7	185.5	7.4	280	3	US-09-029-603-2
8	184.5	7.3	249	4	US-09-134-000C-6713
9	184.5	7.3	318	4	US-09-382-906A-2
10	180.5	7.2	275	4	US-09-266-965-110
11	177	7.0	115	4	US-09-266-965-6
12	173.5	6.9	348	4	US-09-118-637A-4
13	173	6.9	359	4	US-09-328-352-8005
14	172	6.8	275	3	US-09-036-987A-7
15	172	6.8	275	3	US-09-370-700-7
16	172	6.8	275	4	US-09-603-207-7
17	169.5	6.7	114	4	US-09-266-965-2
18	167.5	6.7	361	4	US-09-439-554-30
19	165	6.6	256	4	US-09-328-352-7702
20	163.5	6.5	361	3	US-09-041-718-4
21	161.5	6.4	382	4	US-09-644-907B-11
22	161	6.4	115	4	US-09-266-965-1
23	158	6.3	327	4	US-09-352-991A-28744
24	156.5	6.2	258	4	US-09-134-001C-3244
25	155.5	6.2	385	4	US-09-644-907B-2
26	154.5	6.1	295	4	US-09-540-236-3434
27	154.5	6.1	317	4	US-09-489-039A-8044

28	154.5	6.1	376	3	US-09-041-718-2	Sequence 2, Appli
29	153	6.1	383	3	US-09-041-718-3	Sequence 3, Appli
30	151	6.0	390	4	US-09-489-039A-9705	Sequence 9705, Ap
31	150	6.0	282	1	US-08-457-245-9	Sequence 9, Appli
32	149.5	5.9	116	4	US-09-266-965-7	Sequence 7, Appli
33	149	5.9	253	4	US-09-489-039A-8084	Sequence 8084, Ap
34	145.5	5.8	407	4	US-09-543-681A-4434	Sequence 4434, Ap
35	144.5	5.7	404	4	US-09-252-991A-19166	Sequence 19166, A
36	143.5	5.7	255	4	US-09-543-681A-5713	Sequence 5713, Ap
37	134	5.3	439	4	US-08-311-731A-48	Sequence 48, Appli
38	131.5	5.2	260	4	US-09-489-039A-12651	Sequence 12651, A
39	129	5.1	271	4	US-09-252-991A-21522	Sequence 21522, A
40	126.5	5.0	298	4	US-09-644-907B-8	Sequence 8, Appli
41	124	4.9	209	4	US-09-252-991A-32577	Sequence 32577, A
42	124	4.9	456	4	US-09-252-991A-25142	Sequence 25142, A
43	123	4.9	393	4	US-09-634-238-274	Sequence 274, App
44	122	4.9	253	4	US-09-180-109A-35	Sequence 35, Appli
45	122	4.9	254	4	US-09-180-109A-31	Sequence 31, Appli

#### ALIGNMENTS

RESULT 1  
US-09-266-965-109  
; Sequence 109, Application US/09266965  
; Patent No. 6495348  
; GENERAL INFORMATION:  
; APPLICANT: Sherman, D  
; APPLICANT: Mao, Y  
; APPLICANT: Varoglu, M  
; APPLICANT: He, M  
; APPLICANT: Sheldon, P  
; TITLE OF INVENTION: Mitomycin biosynthetic gene cluster  
; FILE REFERENCE: 600.456US1  
; CURRENT APPLICATION NUMBER: US/09/266,965  
; CURRENT FILING DATE: 1999-03-12  
; EARLIER APPLICATION NUMBER: US 08/624,447  
; EARLIER FILING DATE: 1996-08-19  
; EARLIER APPLICATION NUMBER: PCT/US94/11279  
; EARLIER FILING DATE: 1994-10-06  
; EARLIER APPLICATION NUMBER: US 08/133,963  
; EARLIER FILING DATE: 1993-10-07  
; NUMBER OF SEQ ID NOS: 145  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 109  
; LENGTH: 283  
; TYPE: PRT  
; ORGANISM: Streptomyces lavendulae  
US-09-266-965-109

Query Match 8.7%; Score 218.5; DB 4; Length 283;  
Best Local Similarity 26.9%; Pred. No. 3.7e-15;  
Matches 68; Conservative 46; Mismatches 92; Indels 47; Gaps 11;  
QY 236 YRVFGGVYSTG-----GY-----ETTKEFVSM-----IDLKPGQKVLVDVCGI 275  
DB 23 YDRFTALGAASLGNLHFGYWDSPDSQVPLAATDRLTDMMAERLRCAGSRVLDLGGV 82  
QY 276 GGGDFYMAETFEVVGFDLSVNM1--SEALERSIGLACNEFFVADCTKINYPDNSFDV 333  
DB 83 GTPGVRRLSGARVTGIVSHEQVVRANALAEAGLADRARFQRADANDLPFEDESFDA 142  
QY 334 IYSRDTLHQDPALFRSFYKLPKGGKVLISDYCKKAGPPSPFAFYIKQGYDLDH- 392  
DB 143 VIALESITHMPDRAQLAQVGRVLRPGRLVLTDFPERA-PLAPEGRAAVQRY---LHDF 198  
QY 393 -----VKEYGOMLKQAG-----FVDVLAEDTROPFIVRLKELETVEKEKVFISDFSE 441  
DB 139 MMTWVISAEPPLRGGLWLEEFLLI-SQDTLEKTKFLLS-----ERINSSKQRIETQGE 254  
QY 442 E-----DYNDIVG 449

DB 255 EMVNPDPGDLVG 267

RESULT 2

US-09-041-718-5

Sequence 5, Application US/09041718A

Patent No. 6225075

GENERAL INFORMATION:

APPLICANT: Bard, Martin

TITLE OF INVENTION: DNA encoding sterol methyltransferase

FILE REFERENCE: 740.003051

CURRENT APPLICATION NUMBER: US/09/041.718A

CURRENT FILING DATE: 1998-03-13

NUMBER OF SEQ ID NOS: 9

SOFTWARE: FactSeq for Windows Version 4.0

SEQ ID NO 5

LENGTH: 363

TYPE: PRT

ORGANISM: Triticum atrivum

US-09-041-718-5

Query Match 8.6%; Score 217.5; DB 3; Length 363;

Best Local Similarity 29.8%; Pred. No. 7.2e-15;

Matches 71; Conservative 34; Mismatches 90; Indels 43; Gaps 8;

QY 254 KPSVSM-LDLKPGQKVLVGGCGGDPYMAETFDVGVGFDLSVNMIS--FALERSIGL 310

DB 108 EHFALQLKELKPGMKVLDVGGCGGGLREIARFSSTVTGLANNVDYQITRGKALNRVGL 167

QY 311 KCATFEVADCTKINYPNSPDVYSRDTLHIDQKPAKFRFYKWLKPGKVLISDYC- 369

DB 168 GATCDFVADFMKPPSPDNTDAVVAEATCHAPDPVGCYKRIYVRLKPGQCFAYVNCI 227

QY 370 KAGSPSPFAAYIK--ORGVLDHVDKEYGQML--KDAAGFVDV---LAEDR----- 413

DB 228 THYDNNATHKRIKDEILGNLPDIRSTQCLOAVKAGFEVVDKLAEDSLPWYL 287

QY 414 -----TEQFIRVLKELTVEKVDVIFSDSEEDYNDIVGG 450

DB 288 PLDPSPFSLSPFLTAVTGRIITRNKVKL--EYVGLAPEGSRQVSSFLKKAAGLVEG 343

RESULT 3

US-09-118-637A-2

Sequence 2, Application US/09118637A

Patent No. 6642434

GENERAL INFORMATION:

APPLICANT: DellaFenna, Dean

APPLICANT: Shintani, David K.

TITLE OF INVENTION: TRANSGENIC PLANTS WITH TOCOPHEROL

TITLE OF INVENTION: METHYLTRANSFERASE

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Charles & Brady

STREET: 1 South Pinckney Street

CITY: Madison

STATE: WI

COUNTRY: US

ZIP: 53701-2113

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/118.637A

FILING DATE: 1998-03-13

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Seay, Nicholas J.

REGISTRATION NUMBER: 27386

REFERENCE/DOCKET NUMBER: 920905.90024

TELECOMMUNICATION INFORMATION:

TELEPHONE: (317)337-4847

TELEFAX: (317)337-4847

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 283 amino acids

TELEPHONE: 608-251-5000

TELEFAX: 608-251-9166

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 317 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-118-637A-2

Query Match 8.4%; Score 210.5; DB 4; Length 317;

Best Local Similarity 26.4%; Pred. No. 3.4e-14;

Matches 55; Conservative 48; Mismatches 72; Indels 33; Gaps 7;

QY 236 YERVFGP-----GVVSTGGYETTKFVSMIDL-----KPGQKVLVGGCGIGG 277

DB 49 WEDVGEHMHGYYGPHGTVIDRRQADLIKELLAWAVPQNSAKP-RKILDLGCGIGG 107

QY 278 GDFVMAETFDVGVGFDLSVNMISFALE--RSIGLKCAVEFVADCTKINYPNSFDVIY 335

DB 108 SLLTAQHQAEVNGASLSPQVVERAGERARALGLSTCOFQVANALDLPASDSFDVWV 167

QY 336 SRDTLHIDQKPAKFRFYKWLKPGKVLISDYCKA-----GPPSPFAAYIKQGVYDL 390

DB 168 SLEGEHMPNKAQFLQAEWVLPKGGRLILATWCHRPIDPGNGPLTADERRHL-QAIYDV 226

QY 391 H-----DVKEYGQMLKADAGFVDVLAED 412

DB 227 YCLPVVSLPDEYAEARECGGGEIKTAD 254

RESULT 4

US-09-036-987A-13

Sequence 13, Application US/09036987A

Patent No. 6143526

GENERAL INFORMATION:

APPLICANT: Baltz, Richard H.

APPLICANT: Broughton, Mary C.

APPLICANT: Crawford, Kathryn P.

APPLICANT: Madduri, Krishnamurthy

APPLICANT: Merlo, Donald J.

APPLICANT: Treadway, Patti J.

APPLICANT: Turner, Jan R.

TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide

TITLE OF INVENTION: Production

NUMBER OF SEQUENCES: 39

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dow Agrosciences LLC Patent Department

STREET: 9330 Zionsville Road

CITY: Indianapolis

STATE: Indiana

COUNTRY: USA

ZIP: 46268

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM-PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/036.987A

FILING DATE: 09-MAR-1998

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Stuart, Donald R.

REGISTRATION NUMBER: 28,479

REFERENCE/DOCKET NUMBER: 50,608

TELECOMMUNICATION INFORMATION:

TELEPHONE: (317)337-4816

TELEFAX: (317)337-4847

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 283 amino acids

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-036-987A-13

Query Match  
Best Local Similarity 7.7%; Score 194.5; DB 3; Length 283;  
Best Local Similarity 24.1%; Pred. No. 1.6e-12;  
Matches 57; Conservative 47; Mismatches 106; Indels 27; Gaps 6;  
234 LAYERVFGPGVSTGGY-----ETTKFVSMCLKPGQKVLDPGCGIGGDD 279  
Db 8 LAHGRPLHGYWA--GGYREDAGATPMSDAADQLTDLFIDKAAALRPGAHFLDLGCGGQPV 66  
QY 280 FYMAETFEVVGFDLSVNMISFA--LERSIGLKCAVEFEVADCTKINYPDINSFVYSR 337  
Db 67 VRAACASGVRVTGITVNAQHLAATRLANETGLAGSLFPLVDGAGLPYDFGFFQAAWAM 126  
QY 338 DTILHIQDKPALFRFYKWLKPGKVLISDYCKKAGPPSPPEFAAYIKORGYDLHDVKEYG 397  
Db 127 QSVVQIVDQAAAIREVHRLPEGRFVLGDIITRVRLPE--EYAAV--WTGTTAHTLNSFT 183  
QY 398 QMLKADGFDVLAEDRTEQ-----FIRVLRKELETVEKEKDVFSDFSEEDYNDI 447  
Db 184 ALVSEAGFEILEVTDLTATQTCMVSWYVDLLAKLDELAGVPEAAVGTQQRYLGD 240

## RESULT 5

US-09-370-700-13  
Sequence 13, Application US/09370700  
Patent No. 6274350

GENERAL INFORMATION:  
APPLICANT: Baltz, Richard H  
APPLICANT: Broughton, Mary C  
APPLICANT: Crawford, Kathryn P  
APPLICANT: Madduri, Krishnamurthy  
APPLICANT: Treadway, Patti J  
APPLICANT: Turner, Jan R  
APPLICANT: Waldron, Clive  
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide  
CURRENT APPLICATION NUMBER: US/09/370,700  
EARLIER FILING DATE: 1999-08-09  
EARLIER FILING DATE: 1998-03-09  
NUMBER OF SEQ ID NOS: 39  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 13  
LENGTH: 283  
TYPE: PRT  
ORGANISM: Saccharopolyspora spinosa

US-09-370-700-13  
Query Match  
Best Local Similarity 7.7%; Score 194.5; DB 3; Length 283;  
Matches 57; Conservative 47; Mismatches 106; Indels 27; Gaps 6;

234 LAYERVFGPGVSTGGY-----ETTKFVSMCLKPGQKVLDPGCGIGGDD 279  
Db 8 LAHGRPLHGYWA--GGYREDAGATPMSDAADQLTDLFIDKAAALRPGAHFLDLGCGGQPV 66  
QY 280 FYMAETFEVVGFDLSVNMISFA--LERSIGLKCAVEFEVADCTKINYPDINSFVYSR 337  
Db 67 VRAACASGVRVTGITVNAQHLAATRLANETGLAGSLFPLVDGAGLPYDFGFFQAAWAM 126  
QY 338 DTILHIQDKPALFRFYKWLKPGKVLISDYCKKAGPPSPPEFAAYIKORGYDLHDVKEYG 397  
Db 127 QSVVQIVDQAAAIREVHRLPEGRFVLGDIITRVRLPE--EYAAV--WTGTTAHTLNSFT 183  
QY 398 QMLKADGFDVLAEDRTEQ-----FIRVLRKELETVEKEKDVFSDFSEEDYNDI 447  
Db 184 ALVSEAGFEILEVTDLTATQTCMVSWYVDLLAKLDELAGVPEAAVGTQQRYLGD 240

## RESULT 6

US-09-603-207-13  
Sequence 13, Application US/09603207B  
Patent No. 6521406

GENERAL INFORMATION:  
APPLICANT: Baltz, Richard H  
APPLICANT: Broughton, Mary C  
APPLICANT: Crawford, Kathryn P  
APPLICANT: Madduri, Krishnamurthy  
APPLICANT: Treadway, Patti J  
APPLICANT: Turner, Jan R  
APPLICANT: Waldron, Clive  
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide  
CURRENT APPLICATION NUMBER: US/09/603,207B  
EARLIER FILING DATE: 2000-06-23  
EARLIER FILING DATE: 1998-03-09  
NUMBER OF SEQ ID NOS: 39  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 13  
LENGTH: 283  
TYPE: PRT  
ORGANISM: Saccharopolyspora spinosa

## US-09-603-207-13

Query Match  
Best Local Similarity 7.7%; Score 194.5; DB 4; Length 283;  
Matches 57; Conservative 47; Mismatches 106; Indels 27; Gaps 6;

234 LAYERVFGPGVSTGGY-----ETTKFVSMCLKPGQKVLDPGCGIGGDD 279  
Db 8 LAHGRPLHGYWA--GGYREDAGATPMSDAADQLTDLFIDKAAALRPGAHFLDLGCGGQPV 66  
QY 280 FYMAETFEVVGFDLSVNMISFA--LERSIGLKCAVEFEVADCTKINYPDINSFVYSR 337  
Db 67 VRAACASGVRVTGITVNAQHLAATRLANETGLAGSLFPLVDGAGLPYDFGFFQAAWAM 126  
QY 338 DTILHIQDKPALFRFYKWLKPGKVLISDYCKKAGPPSPPEFAAYIKORGYDLHDVKEYG 397  
Db 127 QSVVQIVDQAAAIREVHRLPEGRFVLGDIITRVRLPE--EYAAV--WTGTTAHTLNSFT 183  
QY 398 QMLKADGFDVLAEDRTEQ-----FIRVLRKELETVEKEKDVFSDFSEEDYNDI 447  
Db 184 ALVSEAGFEILEVTDLTATQTCMVSWYVDLLAKLDELAGVPEAAVGTQQRYLGD 240

## RESULT 7

US-09-029-603-2  
Sequence 2, Application US/09029603  
Patent No. 6210935

GENERAL INFORMATION:  
APPLICANT: Schnupp, Thomas  
APPLICANT: Engel, Natalie  
APPLICANT: Bietenhader, Jurg  
APPLICANT: Toupet, Christine  
APPLICANT: Pospiech, Andreas  
TITLE OF INVENTION: Staurosporin Biosynthesis Gene Clusters  
FILE REFERENCE: 4-20555/A/PCT  
CURRENT APPLICATION NUMBER: US/09/029,603  
EARLIER FILING DATE: 1998-03-20  
EARLIER FILING DATE: 1998-08-19  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2  
LENGTH: 280  
TYPE: PRT  
ORGANISM: Streptomyces longisporoflavus

## FEATURES:

OTHER INFORMATION: methyl transferase-like protein

US-09-029-603-2



```

Query Match      7.4%; Score 185.5; DB 3; Length 280;
Best Local Similarity 26.1%; Pred. No. 1.5e-11;
Matches 68; Conservative 42; Mismatches 92; Indels 59; Gaps 13;

QY 236 YERY-----FGP-----GYSTGGYETIKK-----FVSMDLKPGQKVLNVC 273
DB 18 YDLTSLAMNDGSPNVHIGYWDTPGSEATIEAMDRLTDVFIERLNAYATSHVLDGC 77
QY 274 GIGGDFMAETDFEVGFDLSVNMISPA--LERSIGLKCAVEPEVADCTKINYPNSF 331
DB 78 GVGGPGLRVARTCARVTGISSEQIRPANLAAEAGVADRAVQHGDKMLPFADASP 137
QY 332 DVIVSRDTHIQKALPFRSFKWMLKPGKVLISDYCKKAGPPSPFAANYIKORYD-- 389
DB 138 DAVMALESICHMPDROQVTEVCVRVLRPGGRIVLTDIFER-----HPRKA--VRHFGIDKF 191
QY 390 -----LHDVKEYGQMLKDG-----FVTLAEDRTEQFIRVLKR--LETVEKQKV 434
DB 192 CRDLMTTADIDVALLHRSGLRLRIVDV-----TEQTLRLADEIGRLAAVE-ERP 245
QY 435 FLSQ-----FSEEDY--NDIVG 449
DB 246 ANDGNGFAGDDSPFSLAG 266

```

```

RESULT 8
US-09-134-000C-6713
; Sequence 6713, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134.000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 6713
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-6713

Query Match      7.3%; Score 184.5; DB 4; Length 249;
Best Local Similarity 25.6%; Pred. No. 1.6e-11;
Matches 72; Conservative 49; Mismatches 109; Indels 51; Gaps 12;

QY 209 QKVDSKDDK-----GQRFPLDTQYKNSILRYRVPVPGVSTGGYETTKKPFVSMDLK 263
DB 2 RRVQMKENYKDDNIFQKYQSRSQ-----KGLAGAWETLAKWLP--DPK 47
QY 264 PQKVLNVCIGGIGGGDFMAETDFEVGFDLSVNMISPALERSIGLKCAVEPEVADCTK 323
DB 48 -GKRVLDLGGYGMWCIYAMENGASSVVGVDISHKMLFVAKGKTHFPQ--IETECCAIED 104
QY 324 INTPDMSFVIVSRDTHIQKALPFRSFKWMLKPGKVLISDYCKKAGPPSPFAA-- 381
DB 105 VDPFESFVILLSSLAFFHVAVDYENLKKIYRMLKAGGNLFT-----VEHFPVTANG 157
QY 382 ----YIKQCYDLHDVKE--YQMLKADGAFVDVLAEDRTEQFIRVLKELTEVEKQKV- 434
DB 158 TDQWYNEKEILLHPVDVNYEGKRTAMFL-----EKKYKHTLTYLNTLLNSFII 213
QY 435 --FISDFSRBYNDIVGWNDKLRTAKGQRMGLFWAKK 473
DB 214 NQIVEPQPPENMDI-PGMADEKRRPM-----LIVSACK 247

```

```

RESULT 9
US-09-382-906A-2

```

```

; Sequence 2, Application US/09382906A
; Patent No. 6448475
; GENERAL INFORMATION:
; APPLICANT: Dellapenna, Dean
; APPLICANT: Sinitani, David
; TITLE OF INVENTION: Manipulation of Tocopherol Levels in Transgenic Plants
; FILE REFERENCE: 920905.90032
; CURRENT APPLICATION NUMBER: US/09/382,906A
; CURRENT FILING DATE: 1999-08-25
; PRIOR FILING DATE: 1999-08-25
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Synecocystis PCC6803
US-09-382-906A-2

```

```

Query Match      7.3%; Score 184.5; DB 4; Length 318;
Best Local Similarity 29.6%; Pred. No. 2.3e-11;
Matches 47; Conservative 30; Mismatches 61; Indels 21; Gaps 5;

QY 227 QYKNSILRYRVPVPGVSTGGY--ETTKFV-SMLD-----LKPQKVL 269
DB 42 QWTEGILEY---YMGDHIHLGHYGGPPVAKQFIQSKIDFVHMAQWGLDLPFGTTVL 98
QY 270 DVGGIGGGDFMAETDFEVGFDLSVNMISPALERSIGLKCAVEPEVADCTKINYPDN 329
DB 99 DVGGIGGGSSRLAKDYGFNTGTTISPOQVGRATELT--PPDVTAKFAVDAMALSPFDG 157
QY 330 SPDVIVSRDTHIQKALPFRSFKWMLKPGKVLISDY 368
DB 158 SPDVVVSVEAGPHFMDKAVFAKELLRVVKGGLVVDW 196

```

```

RESULT 10
US-09-266-965-110
; Sequence 110, Application US/09266965
; Patent No. 6495348
; GENERAL INFORMATION:
; APPLICANT: Sherman, D
; APPLICANT: Mao, Y
; APPLICANT: Varoglu, M
; APPLICANT: He, M
; APPLICANT: Sheldon, P
; TITLE OF INVENTION: Mitomycin biosynthetic gene cluster
; FILE REFERENCE: 600.45GUSI
; CURRENT APPLICATION NUMBER: US/09/266,965
; CURRENT FILING DATE: 1999-03-12
; EARLIER FILING DATE: US 08/624,447
; EARLIER FILING DATE: 1996-08-19
; EARLIER FILING DATE: 1994-10-06
; EARLIER FILING DATE: 1994-10-06
; EARLIER FILING DATE: 1993-10-07
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 110
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Streptomyces lavendulae
US-09-266-965-110

```

```

Query Match      7.2%; Score 180.5; DB 4; Length 275;
Best Local Similarity 26.2%; Pred. No. 5e-11;
Matches 59; Conservative 39; Mismatches 105; Indels 19; Gaps 7;

QY 242 PGVSTGGYETTKKPFVSMDLKPGKVLNVCIGGIGGGDFMAETDFEVGFDLSVNMIS 301
DB 41 PTVBEAGRLTDVSRRLRPAQGERVLDVSSGKATLRAARHVRATG--VSNPYQ 98
QY 302 PALERSIGLK--CAVEPEVADCTKINYPDMSFDVIVSRDTHIQKALPFRSFKWMLK 358

```

99	V	L	S	H	L	A	E	G	B	E	T	E	R	I	G	M	L	A	F	P	G	S	F	D	A	C	A	I	E	S	C	H	A	L	E	A	D	V	F	T	E	A	R	V	L	R	:	158						
359	P	G	G	V	L	I	S	Y	C	K	A	G	P	P	S	P	E	F	A	I	K	I	G	Y	D	L	H	D	V	K	E	Y	G	M	L	K	D	A	G	F	V	D	V	L	A	E	D	R	T	E	-	Q	416	
159	P	G	R	V	T	D	P	V	L	R	-	P	L	S	D	A	S	R	T	I	V	D	T	A	-	-	N	D	N	F	Q	Q	F	P	V	L	T	R	E	A	E	D	C	H	R	S	V	G	L	E	V	W	E	213
417	F	I	R	V	-	-	L	R	K	B	E	T	E	K	E	-	-	-	-	K	D	V	F	I	S	D	F	E	E	D	N	I	D	I	V	G	450																	
214	F	L	D	I	G	E	V	P	S	E	A	V	A	K	M	A	R	A	R	D	E	L	G	S	H	M	D	E	A	F	H	R	M	V	D	G	254																	

```

RESULT 11
US-09-266-965-6
: Sequence 6, Application US/09266965
: Patent No. 6495348
: GENERAL INFORMATION:
: APPLICANT: Sherman, D
: APPLICANT: Mao, Y
: APPLICANT: Varoglu, M
: APPLICANT: He, M
: APPLICANT: Sheldon, P
: TITLE OF INVENTION: Mitomycin biosynthetic gene cluster
: FILE REFERENCE: 600.456US1
: CURRENT APPLICATION NUMBER: US/09/266,965
: CURRENT FILING DATE: 1999-03-12
: EARLIER APPLICATION NUMBER: US 08/624,447
: EARLIER FILING DATE: 1996-08-19
: EARLIER APPLICATION NUMBER: PCT/US94/11279
: EARLIER FILING DATE: 1994-10-06
: EARLIER APPLICATION NUMBER: US 08/133,963
: EARLIER FILING DATE: 1993-10-07
: NUMBER OF SEQ ID NOS: 145
: SOFTWARE: PastSeq for Windows Version 3.0
: SEQ ID NO 6
: LENGTH: 115
: TYPE: PRT
: ORGANISM: Amycolatopsis mediterranei
US-09-266-965-6

```

```

RESULT 12
US-09-118-637A-4
/ Sequence 4, Application US/09118637A
/ Patent No. 6642434
/ GENERAL INFORMATION:
/ APPLICANT: DellaPenna, Dean
/ APPLICANT: Shintani, David K.
/ TITLE OF INVENTION: TRANSGENIC P
/ TITLE OF INVENTION: METHYLTRANSF
/ NUMBER OF SEQUENCES: 10
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Quarles & Brady
/ STREET: 1 South Pinckney Street
/ CITY: Madison
/ STATE: WI
/ COUNTRY: US
/ ZIP: 53701-2113
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible

```

```

/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/118,637A
/ FILING DATE:
/ CLASSIFICATION: 800
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Seay, Nicholas J.
/ REGISTRATION NUMBER: 27386
/ REFERENCE/DOCKET NUMBER: 920905.90024
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 608-251-5000
/ TELEFAX: 608-251-9166
/ INFORMATION FOR SEQ ID NO: 4:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 348 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-09-118-637A-4

```

Query Match	6.9%;	Score 173.5;	DB 4;	Length 348;
Best Local Similarity	27.2%;	Pred. No. 4.3e-10;		
Matches 52;	Conservative 30;	Mismatches 64;	Indels 45;	Gaps 7;
Qy	202	NOISLWLNKV-----DSK---DDKGFQ-----RFLDTQYKNSILRYVEFGP	242	
Db	69	NETSLGMEIWDNRHGHGFDYDPOSSVOLSDSGHKEAQLRMIEES-----LRFAGV---	118	
Qy	243	GYVSTGGYETTKFVFVSMLDLKPQKVLDPVCGGIGGGDFYMAETFDVVVGGFDLSVNNMISF	302	
Db	119	-----TDEE-----EEKTKKVVDCVCGIGSSRYLASKPQAGCIGITLSPVQAKR	164	
Qy	303	A--LERSIGLKCAVEFVADCTKINYDPNSFDVIYSRDTILHIQDKPALFYSFYKMLKPG	360	
Db	165	ANDLAAQSLSHKASQVADALDQFEDGKGFDLVWSMESGEHMPDKAFVKVELVRAAPG	224	
Qy	361	GKVLISDYCKK	371	
Db	225	GRIIIVTWCHR	235	

```

RESULT 13
US-09-328-352-8005
; Sequence 8005, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Briston et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO
; ACID SEQUENCES FOR THE DIAGNOSIS OF
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSIS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328.352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 8005
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-8005

```

Query Match	6.9%; Score 173; DB 4; Length 359;
Best Local Similarity	25.6%; Pred. No. 5.2e-10;
Matches	69; Conservative 48; Mismatches 15; Indels 58; Gaps 12
QY	195 VRKKQHONQISWLKQVDSKDDKGFQRFLDTSQVKNSILRLRVPGVYVTGGVETTK 254
DB	124 VRTEDRAQAVSFHVSASKYD-----LNDLLM-----SFGHRLRWK 160
QY	255 EF-VSMLDLKPPQKVQLDVGCIGGGDFYMAETFDVE-----VVGPDLNVNMISFALERS 307
DB	161 RFAINMSGVRGQHVLDIAGTGD-----LAKVFSEVGPGQHVVLSDINESMLNVAERDL 216
OY	308 IGLKCA-VEFEADVADCTKIN-VPDMSFDVIVSRDTILHIQDPALPFRSFYKWLKPQGKVLII 365

Db 217 IDAGTNDVFLANAETLEPADNSFDLVITISGLRNVTDKDAALASMFVRLKPGGLLV 276  
Qy 366 SDYCKAGPSPPAAYIKQGVLDHDKVEYQMLK---DAGFVDVLAED-RTEQFIRV 420  
Db 277 LEFSK-----PVTEPFSKL--YDLSPTALPIMGKLVADESSEYKYLAEISIRHPDORT 328  
Qy 421 LKLEETVEKEKQVFIISDFSEEDYNDIVGG 450  
Db 329 LKGMEN-----AGFONCDYHNLITGG 349

## RESULT 14

US-09-036-987A-7  
; Sequence 7, Application US/09036987A  
; Patent No. 6143526  
; GENERAL INFORMATION:  
; APPLICANT: Baltz, Richard H.  
; APPLICANT: Broughton, Mary C.  
; APPLICANT: Crawford, Kathryn P.  
; APPLICANT: Madduri, Krishnamurthy  
; APPLICANT: Merlo, Donald J.  
; APPLICANT: Treadway, Patti J.  
; APPLICANT: Turner, Jan R.  
; APPLICANT: Waldron, Clive  
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide  
; NUMBER OF SEQUENCES: 39  
; CORRESPONDENCE ADDRESS:  
; ADDRESSES: Dow Agrosciences LLC Patent Department  
; STREET: 9330 Zionsville Road  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: USA  
; ZIP: 46268

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/036,987A  
; FILING DATE: 09-MAR-1998  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stuart, Donald R.  
; REGISTRATION NUMBER: 28,479  
; REFERENCE/DOCKET NUMBER: 50,608  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (317)337-4816  
; TELEFAX: (317)337-4847  
; INFORMATION FOR SEQ ID NO: 7:  
; LENGTH: 275 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-036-987A-7

Query Match 6.8%; Score 172; DB 3; Length 275;  
Best Local Similarity 23.3%; Pred. No. 4.3e-10;  
Matches 55; Conservative 39; Mismatches 102; Indels 40; Gaps 6;  
Qy 241 GP-----GYVSTGGYETTKFVSMLD-----LKPQKVLVDGCGIGGGDFYMAETFD 287  
Db 29 GPCAIHHGYWENDGRASWQQAADRLDLVAERTVLDGCVRLDVGCGTGQPALRVADNA 88  
Qy 288 VEVVGFGLSVNMISFALR--RSIGLKCAVEFVADCTKINYPDNSFDVIYSRDTLHIQD 345  
Db 89 IQTGITVSGVQVAIAADCAERGLSHRVDFSCVDAMSIFYPDNAPDAAMQSLLENSE 148  
Qy 346 KPALFRSFYKWLKPGKVLISDYCKKAGPPSPPAAYIKQGVLDHDKVEYQMLKADGAF 405

Db 149 PDRAIREILRVLPKGGILGVTEVYVYKREAGGMPVSGDRMPTGLRICLAELQLLESRAAGF 208  
Qy 406 VDVLAEORTQFIRVLKKELETVEKEKQVFIISDFSEE-----DYNDIVGGW 451  
Db 209 -EIL-----DWEDEVSSRTRVFMFPQFAEELAAHQHGIADRYGPAVAGM 249  
RESULT 15  
US-09-370-700-7  
; Sequence 7, Application US/09370700  
; Patent No. 6274350  
; GENERAL INFORMATION:  
; APPLICANT: Baltz, Richard H.  
; APPLICANT: Broughton, Mary C.  
; APPLICANT: Crawford, Kathryn P.  
; APPLICANT: Madduri, Krishnamurthy  
; APPLICANT: Treadway, Patti J.  
; APPLICANT: Turner, Jan R.  
; APPLICANT: Waldron, Clive  
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide  
; FILE REFERENCE: 50489 DIV1  
; CURRENT APPLICATION NUMBER: US/09/370,700  
; CURRENT FILING DATE: 1999-08-09  
; EARLIER APPLICATION NUMBER: US 09/36987  
; EARLIER FILING DATE: 1998-03-09  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 275  
; TYPE: PNT  
; ORGANISM: Saccharopolyspora spinosa  
US-09-370-700-7

Query Match 6.8%; Score 172; DB 3; Length 275;  
Best Local Similarity 23.3%; Pred. No. 4.3e-10;  
Matches 55; Conservative 39; Mismatches 102; Indels 40; Gaps 6;  
Qy 241 GP-----GYVSTGGYETTKFVSMLD-----LKPQKVLVDGCGIGGGDFYMAETFD 287  
Db 29 GPCAIHHGYWENDGRASWQQAADRLDLVAERTVLDGCVRLDVGCGTGQPALRVADNA 88  
Qy 288 VEVVGFGLSVNMISFALR--RSIGLKCAVEFVADCTKINYPDNSFDVIYSRDTLHIQD 345  
Db 89 IQTGITVSGVQVAIAADCAERGLSHRVDFSCVDAMSIFYPDNAPDAAMQSLLENSE 148  
Qy 346 KPALFRSFYKWLKPGKVLISDYCKKAGPPSPPAAYIKQGVLDHDKVEYQMLKADGAF 405  
Db 149 PDRAIREILRVLPKGGILGVTEVYVYKREAGGMPVSGDRMPTGLRICLAELQLLESRAAGF 208  
Qy 406 VDVLAEORTQFIRVLKKELETVEKEKQVFIISDFSEE-----DYNDIVGGW 451  
Db 209 -EIL-----DWEDEVSSRTRVFMFPQFAEELAAHQHGIADRYGPAVAGM 249

Search completed: July 26, 2004, 13:16:23  
Job time : 20 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2004, 13:15:32 ; Search time 46 Seconds  
(without alignments)  
3220.042 Million cell updates/sec

Title: US-10-031-331B-40

Perfect score: 2515

Sequence: 1 HVDTITAMMLDSQASDL.....KLRTAKGEQRWGLFVAKK 473

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1288442 seqs, 313154207 residues

Total number of hits satisfying chosen parameters: 1288442

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2011	80.0	513	12	US-10-425-114-37470 Sequence 37470, A
2	1932.5	76.8	530	16	US-10-437-963-155711 Sequence 155711, A
3	1829.5	72.7	491	14	US-10-149-759-82 Sequence 82, Appl
4	1760	70.0	501	12	US-10-424-599-172601 Sequence 172601, A
5	1654	65.8	399	12	US-10-425-114-72568 Sequence 72568, A
6	1321	52.5	314	12	US-10-425-114-69972 Sequence 69972, A
7	1086	43.3	331	12	US-10-425-114-44987 Sequence 44987, A
8	1064	42.3	377	12	US-10-425-114-60338 Sequence 60338, A
9	937	37.3	219	12	US-10-425-114-53753 Sequence 53753, A
10	839	33.4	237	12	US-10-424-599-257495 Sequence 257495, A
11	690.5	27.5	271	16	US-10-437-963-110035 Sequence 110035, A
12	665.5	26.5	183	12	US-10-424-599-240673 Sequence 240673, A
13	658	26.2	168	14	US-10-149-759-50 Sequence 50, Appl
14	516.5	20.5	437	15	US-10-369-493-6306 Sequence 6306, Ap
15	516.5	20.5	437	16	US-10-602-268-21 Sequence 21, Appl

16	496.5	19.7	472	16	US-10-602-268-12	Sequence 12, Appl
17	480	19.1	437	16	US-10-602-268-11	Sequence 11, Appl
18	425	16.9	161	12	US-10-424-599-224095	Sequence 224095, A
19	421	16.7	133	16	US-10-437-963-110034	Sequence 110034, A
20	413	16.4	475	16	US-10-602-268-19	Sequence 19, Appl
21	413	16.4	484	16	US-10-602-268-20	Sequence 20, Appl
22	404	16.1	460	16	US-10-602-268-7	Sequence 7, Appl
23	399	15.9	460	16	US-10-602-268-8	Sequence 8, Appl
24	396	15.7	457	16	US-10-602-268-9	Sequence 9, Appl
25	375	14.9	120	16	US-10-437-963-155679	Sequence 155679, A
26	353.5	14.1	469	16	US-10-602-268-10	Sequence 10, Appl
27	296	11.8	155	12	US-10-424-599-216620	Sequence 216620, A
28	286.5	11.5	264	15	US-10-369-493-12004	Sequence 12004, A
29	245	9.7	111	16	US-10-437-963-149311	Sequence 149311, A
30	243.5	9.7	117	12	US-10-424-599-174519	Sequence 174519, A
31	241.5	9.6	281	14	US-10-132-134-20	Sequence 20, Appl
32	224	8.9	280	14	US-10-279-029-106	Sequence 106, Appl
33	224	8.9	280	14	US-10-219-810-39	Sequence 39, Appl
34	218.5	8.7	283	10	US-09-953-348-109	Sequence 109, Appl
35	218.5	8.7	283	14	US-10-267-255-109	Sequence 109, Appl
36	214	8.5	280	14	US-10-279-029-107	Sequence 107, Appl
37	214	8.5	280	14	US-10-219-810-40	Sequence 40, Appl
38	214	8.5	286	15	US-10-369-493-18938	Sequence 18938, A
39	210.5	8.4	317	14	US-10-219-810-41	Sequence 41, Appl
40	201.5	8.0	243	15	US-10-369-493-3806	Sequence 3806, Ap
41	201.5	8.0	362	16	US-10-437-963-147768	Sequence 147768, A
42	200.5	8.0	345	14	US-10-279-029-95	Sequence 95, Appl
43	200.5	8.0	345	14	US-10-219-810-22	Sequence 22, Appl
44	199	7.9	344	9	US-09-779-144-7	Sequence 7, Appl
45	198	7.9	338	12	US-10-425-114-41948	Sequence 41948, A

#### ALIGNMENTS

#### RESULT 1

US-10-425-114-37470  
; Sequence 37470, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 37470  
; LENGTH: 513  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURES:  
; OTHER INFORMATION: Clone ID: LIB22-074-F3\_FLI.pep  
US-10-425-114-37470

Query Match	80.0%	Score	2011	DB	12	Length	513
Best Local Similarity	77.8%	Pred. No.	9.4e-189				
Matches	367	Conservative	51	Mismatches	54	Indels	0
						Gaps	0
QY	1	HVDTITAMMLDSQASDLKKEPPEILSMPLPEKCKLLELGAGIGRTGELAKAGQV	60				
Db	41	HSADITVENMLDSRASDLKKEPPEILSLPPEYKSVLELGAGIGRTGELAKAGEL	100				
QY	61	IALDPIEAIKKNVINGHYKNVKFMCAVTSPTLSFPFHSLDVIFSNWLLMYLSDREEV	120				
Db	101	IALDPIFNVIKNSINGHYKNVKFMCAVTSPTLSFPFHSLDVIFSNWLLMYLSDREEV	160				
QY	121	NLVERMLKWLPGGYIFPSCFHSQGDHKKNSNPHYREPPYTKAFKCHLQDGSQNS	180				

Db 161 LLAEWGMKVGQYIFRSCFHQSDSKRKNPHTYRPFYKVFQECOTRDAAGNS 220  
 Qy 181 YELSLSCKICGAVRKNQKQIOWLMQVDSKDDKGPORLDTSOYKCHSILRYERVF 240  
 Db 221 FELSIMGCKICGAVRKNQKQIOWLMQVDSKDDKGPORLDTSOYKCHSILRYERVF 280  
 Qy 241 GPGVSTGGYTTTKEFVSMMLDKPGQKVLVGGGIGGDFYMAETFDVEVVGFDLSVNM 300  
 Db 281 GCGFVSTGGLETTKEFVSMMLDKPGQKVLVGGGIGGDFYMAETFDVEVVGFDLSVNM 340  
 Qy 301 SPALERSIGLCAVEFEVADCTKINYPDNSFDVYISRDITLHIOKPAFLFZYFKWLPK 360  
 Db 341 SPALERAIGLSCSVEFEVADCTKINYPDNSFDVYISRDITLHIOKPAFLFZYFKWLPK 400  
 Qy 361 GKVLISDCKVAGPSPPEPAAYIKORGYDLHDVKEVGMKLDAGFVDVLAEDRTQPIRV 420  
 Db 401 GKVLISDCKVAGPSPPEPAAYIKORGYDLHDVKEVGMKLDAGFVDVLAEDRTQPIRV 460  
 Qy 421 LRKELETVEKDVIFSDPSEEDYNDIVGGNDKLRRTAKGEQRWGLFVAKK 472  
 Db 461 LRKELDVKEKKEKFIISDFSKEDYDDIVGGWKSCLERCASDEQKWLPIANK 512

RESULT 2  
 US-10-437-963-155711  
 ; Sequence 155711, Application US/10437963  
 ; Publication No. US20040123343A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J.  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Wu, Wei  
 ; APPLICANT: Bouharov, Andrey A.  
 ; APPLICANT: Barbazuk, Brad  
 ; APPLICANT: Li, Ping  
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
 ; FILE REFERENCE: 38-21(53221)B  
 ; CURRENT APPLICATION NUMBER: US/10/437,963  
 ; CURRENT FILING DATE: 2003-05-14  
 ; NUMBER OF SEQ ID NOS: 204966  
 ; SEQ ID NO 155711  
 ; LENGTH: 530  
 ; TYPE: PRT  
 ; ORGANISM: Oryza sativa  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_5544C.1.1.pep  
 US-10-437-963-155711

Query Match 76.8%; Score 1932.5; DB 16; Length 530;  
 Best Local Similarity 71.4%; Pred. No. 5.3e-181;  
 Matches 359; Conservative 52; Mismatches 61; Indels 31; Gaps 2;

Qy 1 HTVDTLEAWMLDSQASDLKKEPRLPILSMLPLEGKLELGLAGIGRFTGELAKAGOV 60  
 Db 28 HSKDLTVANMLDSQASDLKKEPRLPILSMLPLEGKLELGLAGIGRFTGELAKAGOV 76  
 Qy 40 LELGAGIGRFTGELAKAGOVIALDQFISAIKKNVINGHYKXKPMCADVTSPTLSPPP 99  
 Db 88 LELGAGIGRFTGELAKAGOVIALDQFISAIKKNVINGHYKXKPMCADVTSPTLSPPP 147  
 Qy 100 HSLDVIPSNMLLSDERVENLVERMLKWLKPGQYIFRSCFHQSDHKKSNPHTYR 159  
 Db 148 NSIDLIFSNMLLSDERVENLVERMLKWLKPGQYIFRSCFHQSDHKKSNPHTYR 207  
 Qy 160 EPRYTKAFTECHLDQSGNSYELSLSCKICGAVRKNQKQIOWLMQVDSKDDKGP 219  
 Db 208 EPRYTKAFTECHLDQSGNSYELSLSCKICGAVRKNQKQIOWLMQVDSKDDKGP 267  
 Qy 220 QRFLDTSQYKCNLSILRYERVFPGYVSTGGY-----ETTKFVSMMLDKPGQKVL 269

Db 268 QRFLDNVOYKASGILRYERVFPGYVSTGGYVCLFFYLSPSTTKTFVDRDLDPKQNVL 327  
 Qy 270 DVGCGIGGDFYMAETFDVEVVGFDLSVNMISFALERSIGLCAVEFEVADCTKINYPDN 329  
 Db 328 DVGCGIGGDFYMAETFDVEVVGFDLSVNMISFALERSIGLCAVEFEVADCTKINYPDN 387  
 Qy 330 SPDVYISRDITLHIOKPAFLFZYFKWLPKGGKVLISDYCKKAGPPSPPEFAAYIKORGYD 389  
 Db 388 TFDVYISRDITLHIOKPAFLFZYFKWLPKGGKVLISDYCKKAGPPSPPEFAAYIKORGYD 447  
 Qy 390 LHDVKEVGMKLDAGFVDVLAEDRTQPIRVLAKELETVEKEDVIFSDPSEEDYNDIVG 449  
 Db 448 LHDVRAVQMLENAGPHDVIABDRTQPIRVLAKELETVEKEDVIFSDPSEEDYNDIVG 507  
 Qy 450 GNDKLRRTAKGEQRWGLFVAKK 472  
 Db 508 GNDKLRRTAKGEQRWGLFVAKK 530

RESULT 3  
 US-10-149-759-82  
 ; Sequence 82, Application US/10149759  
 ; Publication No. US20030157592A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lerchl, Jens  
 ; APPLICANT: Renz, Andreas  
 ; APPLICANT: Ehrhardt, Thomas  
 ; APPLICANT: Reinold, Andreas  
 ; APPLICANT: Cirpus, Petra  
 ; TITLE OF INVENTION: Moss Genes from Physcomitrella patens encoding proteins  
 ; FILE REFERENCE: BASF/NAB 1333/99 PCT/US  
 ; CURRENT APPLICATION NUMBER: US/10/149,759  
 ; CURRENT FILING DATE: 2002-10-17  
 ; PRIOR APPLICATION NUMBER: PCT/EP/00/12698  
 ; PRIOR FILING DATE: 2000-12-14  
 ; NUMBER OF SEQ ID NOS: 82  
 ; SOFTWARE: WordPerfect, version 6.1  
 ; SEQ ID NO 82  
 ; LENGTH: 491  
 ; TYPE: PRT  
 ; ORGANISM: Physcomitrella patens  
 US-10-149-759-82

Query Match 72.7%; Score 1829.5; DB 14; Length 491;  
 Best Local Similarity 70.0%; Pred. No. 6.6e-171;  
 Matches 332; Conservative 65; Mismatches 74; Indels 3; Gaps 2;

Qy 1 HTVDTLEAWMLDSQASDLKKEPRLPILSMLPLEGKLELGLAGIGRFTGELAKAGOV 60  
 Db 17 HSEVPSVEMMLDSQASDLKKEPRLPILSMLPLEGKLELGLAGIGRFTGELAKAGOV 76  
 Qy 61 IALDFTIESAIKKNVINGHYKXKPMCADVTSPTLSPPPISLVDVIFSNMLLSDERVEN 120  
 Db 77 LAMPFMENTIKKEDVNGHYNNIDFKCADVTSPTLSPPPISLVDVIFSNMLLSDERVEN 136  
 Qy 121 NLVERMLKWLKPGQYIFRSCFHQSDHKKSNPHTYRPFYKVFQECOTRDAAGNS 180  
 Db 137 GLASRVNMLRPGYIFRSCFHQSDHKKSNPHTYRPFYKVFQECOTRDAAGNS 195  
 Qy 181 YELSLSCKICGAVRKNQKQIOWLMQVDSKDDKGPORLDTSOYKCHSILRYERVF 238  
 Db 196 FRFEMGCKICGAVRKNQKQIOWLMQVDSKDDKGPORLDTSOYKCHSILRYERVF 255  
 Qy 239 VFGPGYVSTGGYTTTKEFVSMMLDKPGQKVLVGGGIGGDFYMAETFDVEVVGFDLSV 298  
 Db 256 IFGPGYVSTGGYTTTKEFVSMMLDKPGQKVLVGGGIGGDFYMAETFDVEVVGFDLSV 315  
 Qy 299 MISFALERSIGLCAVEFEVADCTKINYPDNSFDVYISRDITLHIOKPAFLFZYFKW 358  
 Db 316 MISFALERSIGLCAVEFEVADCTKINYPDNSFDVYISRDITLHIOKPAFLFZYFKW 375

QY 359 PGKVLISDYCKKAGPPSDEFAAYIKORGYDLHDVKEYGOMLKADGPFVDVLAEDRTQFI 418  
DB 376 PGKVLISDYCRAPQPSAEFAAYIQORGYDLHVSQKYGEMLEADGPFVVEVVAEDRTQFI 435  
QY 419 RVLRKELETVKEKDVFIISDFSEEDYNDIVGWNKDLARTAKGEORWGLFVAKK 472  
DB 436 EVLQRELATTEAGRDQFINDFSEEDYNYIVSGWKSGLKRCNSDEQKWLFIAYK 489

RESULT 4  
US-10-424-599-172601  
; Sequence 172601, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic, David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 172601  
; LENGTH: 501  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(501)  
; OTHER INFORMATION: unsure at all xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_126075C.1.pap  
US-10-424-599-172601

Query Match 70.0%; Score 1760; DB 12; Length 501;  
Best Local Similarity 68.1%; Pred. No. 4.7e-164;  
Matches 323; Conservative 68; Mismatches 81; Indels 2; Gaps 1;

QY 1 HTVDITIAWMLDSOASDLKEERPEITLSMLPLEGKCLLEAGIGRTGTGLAEKAGQV 60  
DB 27 HSADLSVEMWMLDSKAHLDEERPEVLSLPPPEKSGVSLGALGIGRTGELAKAQL 86  
QY 61 IALDIESAIKNEVINGHYKNKPMCADVTSPTLS--PPHSLDVIISNMLMLSLDER 118  
DB 87 LAVDEIDTAIKNETTINGHHNVKFLCADVTSFNNNNVSEGSVOVFSNKLIMYLSID 146  
QY 119 VENLVERMLKXPGGYIFFRESCHFQSGDHKRSNPTHYRPFYTKAFKECHLDQSG 178  
DB 147 VEKFTDRMIRLLIDGGYVFFKGTCTQSGNSKRSKYNSTHYKVPFNTKVIKECHMSDNK 206  
QY 179 NSVELSLSCXKIGAYVNNKQNOISWLAKQVDSKDDKGFORFLDTSQYKNSILRYER 238  
DB 207 NSFELSLGCKKIGAYVNNKQNOICWISLKVRSQDDRGFQRLDRVYSHKSLRYER 266  
QY 239 VFGPGVSTGGYTTKXGPFVSMMLDKPGQKVLDCGCGGDFYMAFTFVGVVGFDSVN 298  
DB 267 MTGPGFVSTGGLETTKTFKFAVGLKPGQKVLDCGCGGDFYMAENFVGVIGIDLSIN 326  
QY 299 MTSFALERSIGLKAVEPADCTKINYPDINSFVIVSRDTHIQOKPALFRSPFVWLK 358  
DB 327 MSLAIERAIGLNAVFECDACVRYTPENTFVIVTRTMLHVKKDTLFRSPFVWLK 386  
QY 359 PGKVLISDYCKKAGPPSDEFAAYIKORGYDLHDVKEYGOMLKADGPFVDVLAEDRTQFI 418  
DB 387 PGKVLISDYCKKAGPPSDEFAAYIKORGYDLHDVKEYGOMLKADGPFVDVLAEDRTQFI 446  
QY 419 RVLRKELETVKEKDVFIISDFSEEDYNDIVGWNKDLARTAKGEORWGLFVAKK 472  
DB 447 NTLQELNALENKODFDIGDFSEEDYNEIVERWAKQTRCASDEQKWLFIAYK 500

RESULT 5  
US-10-425-114-72568  
; Sequence 72568, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 72568  
; LENGTH: 399  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB23-061-D6\_FLI.pap  
US-10-425-114-72568

Query Match 65.8%; Score 1654; DB 12; Length 399;  
Best Local Similarity 77.1%; Pred. No. 9.2e-154;  
Matches 300; Conservative 40; Mismatches 49; Indels 0; Gaps 0;

QY 84 KPMCADVTSPTLSPPHSLDVIISNMLMLYLSDERVENLVERMLKQVLPKGGYIFFRSCF 143  
DB 10 KPMCADVTSPTLSPKIDTGLDLSNMLMLYLSDEKVELLAERWGVWIKVGGYIFFRSCF 69  
QY 144 HQSGDHKRSNPTHYRPFYTKAFKECHLDQSGNSYELSLSCXKIGAYVNNKQNO 203  
DB 70 HQSGDHKRSNPTHYRPFYTKAFKECHLDQSGNSYELSLSCXKIGAYVNNKQNO 129  
QY 204 ISWLAKQVDSKDDKGFORFLDTSQYKNSILRYERFVGVVSTGGYTTKXGPFVSMMLDK 263  
DB 130 ICWIMQKVSSENDGRFRLDNVQKSSGLRYRFFVGGQFVSTGGLETTKEFVEKWLK 189  
QY 264 PQOKVLDVCGGIGGDFYMAETFDVGVGFDLSVNNISPALERSIGLKCAVEFEVADCTK 323  
DB 190 PQOKVLDVCGGIGGDFYMAETFDVGVGFDLSVNNISPALERSIGLKCAVEFEVADCTK 249  
QY 324 INYPDINSFVIVSRDTHIQOKPALFRSPFVWLKPGQKVLISDYCKKAGPPSPFAAYI 383  
DB 250 KHYPDNSFVIVSRDTHIQOKPALFRSPFVWLKPGQKVLISDYCKKAGPPSPFAAYI 309  
QY 384 KORGYDLHDVKEYGOMLKADGPFVDVLAEDRTQFIIRLAKLESTVEKEDVPISDFSEED 443  
DB 310 KORGYDLHDVKEYGOMLKADGPFVDVLAEDRTQFIIRLAKLESTVEKEDVPISDFSEED 369  
QY 444 YNDIVGWNKDLARTAKGEORWGLFVAKK 472  
DB 370 YDDIVGWNKDLARTAKGEORWGLFVAKK 398

RESULT 6  
US-10-425-114-69972  
; Sequence 69972, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114

```

; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59972
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Zea mays subsp. mexicana
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMROTBSINTE108803_FLI.pep
; US-10-425-114-69972

Query Match
Best Local Similarity 52.5%; Score 1321; DB 12; Length 314;
Matches 241; Conservative 33; Mismatches 40; Indels 0; Gaps 0;

Qy 159 RPRFVTKAFKCHLDGSGNSYLSLCKCIGAYVRNKNQNOISLWQKVDKDKG 218
Db 1 RPRFVTKAFKCHLDGSGNSYLSLCKCIGAYVRNKNQNOISLWQKVDKDKG 60
Qy 219 FORPLDTGQYKNSILRYRVFGPGYVSTGGYETTKFVSMCLKPGQKVLVGGGIGGG 278
Db 61 FORFLDNVQYKTSGLIRYRVFGPGYVSTGGYETTKFVSMCLKPGQKVLVGGGIGGG 120
Qy 279 DFVMAETFDVGVGFDLSVNMISFALERSIGLKCAVEFVADCTKINYPDSFVYISRD 338
Db 121 DFVMAETFDVGVGFDLSVNMISFALERSIGLKCAVEFVADCTKINYPDSFVYISRD 180
Qy 339 TILHIQDKPALSFRSFKWLPKGGKVLISDYCKKAGPSPFPAAYIKORGYDLHDVKEYGO 398
Db 181 TILHIQDKPALSFRSFKWLPKGGKVLISDYCKKAGPSPFPAAYIKORGYDLHDVKEYGO 240
Qy 399 MLKADGFDVLAEADRTQPIRVLRKLETVKEKXOVFISDFSEEDYNDIVGGNKLART 458
Db 241 MLKADGFDVLAEADRTQPIRVLRKLETVKEKXOVFISDFSEEDYNDIVGGNKLART 300
Qy 459 AKGEORWGLFVAKK 472
Db 301 SAGEQRWGLFVATK 314

RESULT 7
US-10-425-114-44987
; Sequence 44987, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 44987
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700211781_FLI.pep
; US-10-425-114-44987

Query Match
Best Local Similarity 43.3%; Score 1088; DB 12; Length 331;
Matches 210; Conservative 38; Mismatches 68; Indels 8; Gaps 3;

Qy 153 SNPTH-----YRPRFTKAFKCHLDGSGNSYLSLCKCIGAYVRNKNQNOISLW 208
Db 12 SKPTVLLVLDVYISGNAP-ECNQRKPEGSAL---LNGLAFTKIGIKRPFVIVCWLW 67
Qy 209 QKVDSKDDKRGFORPLDTSQYKNSILRYRVFGPGYVSTGGYETTKFVSMCLKPGQKV 268

; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59972
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Zea mays subsp. mexicana
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMROTBSINTE108803_FLI.pep
; US-10-425-114-69972

Query Match
Best Local Similarity 76.8%; Score 1321; DB 12; Length 314;
Matches 241; Conservative 33; Mismatches 40; Indels 0; Gaps 0;

Qy 159 RPRFVTKAFKCHLDGSGNSYLSLCKCIGAYVRNKNQNOISLWQKVDKDKG 218
Db 1 RPRFVTKAFKCHLDGSGNSYLSLCKCIGAYVRNKNQNOISLWQKVDKDKG 60
Qy 219 FORPLDTGQYKNSILRYRVFGPGYVSTGGYETTKFVSMCLKPGQKVLVGGGIGGG 278
Db 61 FORFLDNVQYKTSGLIRYRVFGPGYVSTGGYETTKFVSMCLKPGQKVLVGGGIGGG 120
Qy 279 DFVMAETFDVGVGFDLSVNMISFALERSIGLKCAVEFVADCTKINYPDSFVYISRD 338
Db 121 DFVMAETFDVGVGFDLSVNMISFALERSIGLKCAVEFVADCTKINYPDSFVYISRD 180
Qy 339 TILHIQDKPALSFRSFKWLPKGGKVLISDYCKKAGPSPFPAAYIKORGYDLHDVKEYGO 398
Db 181 TILHIQDKPALSFRSFKWLPKGGKVLISDYCKKAGPSPFPAAYIKORGYDLHDVKEYGO 240
Qy 399 MLKADGFDVLAEADRTQPIRVLRKLETVKEKXOVFISDFSEEDYNDIVGGNKLART 458
Db 241 MLKADGFDVLAEADRTQPIRVLRKLETVKEKXOVFISDFSEEDYNDIVGGNKLART 300
Qy 459 AKGEORWGLFVAKK 472
Db 301 SAGEQRWGLFVATK 314

RESULT 8
US-10-425-114-60338
; Sequence 60338, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 60338
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3607-019-D6_FLI.pep
; US-10-425-114-60338

Query Match
Best Local Similarity 42.3%; Score 1064; DB 12; Length 377;
Matches 196; Conservative 33; Mismatches 47; Indels 0; Gaps 0;

Qy 1 HTVDLTIRAMLDLSOASDLKKEPEREILSMPLPGLGKLELGAIGRFTGELAERAGOV 60
Db 80 HSGELMLEAIMLDLSRAALDKKEPEREILSMPLPGLGKLELGAIGRFTGELAERAGOV 139
Qy 61 IALDPTIESAIKKNVINGHYKXVKFMCADVTSPFPHSLDVIFSNMLLMYLSDEYVE 120
Db 140 FAVDFVESVIKKNGSINDHYGNTSPMCADVTSPPLMTEANSIDLIFSNMLLMYLSDEY 199
Qy 121 NLVERMLKWLKPGGYVIFRESCFHQSQDHKKSNPHYRPRFTKAFKCHLDGSGNS 180
Db 200 KLVERNKWLKPGGYVIFRESCFHQSQDHKKSNPHYRPRFTKAFKCHLDGSGNS 259
Qy 181 YELSLSCICIGAYVRNKNQNOISLWQKVDKDKGFORPLDTSQYKNSILRYRV 240
Db 260 FKLSLITFCIGAYVRNKNQNOISLWQKVDKDKGFORPLDTSQYKNSILRYRV 319
Qy 241 GPGYVSTGGYETTKFVSMCLKPGQKVLVGGGIGG 276
Db 320 GDGYVSTGGYETTKFVSMCLKPGQKVLVGGGIGG 355

RESULT 9
US-10-425-114-53753
; Sequence 53753, Application US/10425114
```



Publication No. US20040034888A1

GENERAL INFORMATION:

APPLICANT: Liu, Jingdong

APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, Steven E

APPLICANT: Tabaska, Jack E

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53313)B

CURRENT APPLICATION NUMBER: US/10/425.114

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 73128

SEQ ID NO 53753

LENGTH: 219

TYPE: PRT

ORGANISM: Gossypium hirsutum

FEATURE:

OTHER INFORMATION: Clone ID: LIB3829-031-F4\_FLI.pep

US-10-425-114-53753

Query Match 37.3%; Score 937; DB 12; Length 219;

Best Local Similarity 80.7%; Pred. No. 1.5e-83;

Matches 176; Conservative 16; Mismatches 26; Indels 0; Gaps 0;

QY 255 EFVSMDLKPGQKLVDCGCGIGGDFYMAETFDVVEVGFDSLVNMISFALERSIGLKCAV 314

DB 1 EFVAKDLKPGQKLVDCGCGIGGDFYMAKEFGYVVGIDLSINMISFALERANGIDCSV 60

QY 315 EFVADCTKINYPNSFDVYSRDTILHIQDKPALFRSFKYKMLKPGQKLVSDYCKKAGP 374

DB 61 EFVADCTKINYPNSFDVYSRDTILHIQDKPALFRSFKYKMLKPGQKLVSDYCKKAGP 120

QY 375 PSPEFAYIKQGYDLHVKYGGMLKADGPFVDVLAEDRTEQFIRVLRKLETVKKEKDV 434

DB 121 SPSPEFAYIKQGYDLHVKYGGMLKADGPFVDVLAEDRTEQFIRVLRKLETVKKEKDV 180

QY 435 FIDFSEEDYNDVGGNDKLRRTAKGEORWGLFVAKK 472

DB 181 FIDFSEEDYNDVGGNDKLRRTAKGEORWGLFVAKK 218

RESULT 10

US-10-424-599-257495

Sequence 257495, Application US/10424599

Publication No. US20040031072A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J

APPLICANT: Kovalic, David K

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53223)B

CURRENT APPLICATION NUMBER: US/10/424.599

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 285684

SEQ ID NO 257495

LENGTH: 237

TYPE: PRT

ORGANISM: Glycine max

FEATURE:

OTHER INFORMATION: Clone ID: PAT\_MRT3847\_74541C.1.pep

US-10-424-599-257495

Query Match 33.4%; Score 839; DB 12; Length 237;

Best Local Similarity 66.7%; Pred. No. 7.7e-74;

Matches 158; Conservative 30; Mismatches 49; Indels 0; Gaps 0;

QY 237 ERVFGPGYVSTGGYETTKKPFYMLDKPGQKLVDCGCGIGGDFYMAETFDVVEVGFDS 296

DB 1 ESVFGPGYVSTGGYETTKKPFYMLDKPGQKLVDCGCGIGGDFYMAETFDVVEVGFDS 60

QY 297 VNMISFALERSIGLKCAVFEVADCTKINYPNSFDVYSRDTILHIQDKPALFRSFKV 356

DB 61 INIISLAIRAIGLKCCVFEACADCTKCTFPVNTFVYSRDTILHIQDKPALFRSFKV 120

QY 357 LKPGQKLVSDYCKKAGPSPSFAAYIKQGYDLHVKYGGMLKADGPFVDVLAEDRTEQ 416

DB 121 LKRGSTLLITDYCKSEGSLSIGYAEVYKGGYVYHDMKTYCSMLNAGFDDVVAEDRTNL 180

QY 417 FIVLRKLETVKKEKDVFSDFSEEDYNDVGGNDKLRRTAKGEORWGLFVAKK 473

DB 181 FIVLRKLETVKKEKDVFSDFSEEDYNDVGGNDKLRRTAKGEORWGLFVAKK 237

RESULT 11

US-10-437-963-110035

Sequence 110035, Application US/10437963

Publication No. US20040123343A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

APPLICANT: Wu, Wei

APPLICANT: Boukharov, Andrey A.

APPLICANT: Barbazuk, Brad

APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53221)B

CURRENT APPLICATION NUMBER: US/10/437.963

CURRENT FILING DATE: 2003-05-14

NUMBER OF SEQ ID NOS: 204966

SEQ ID NO 110035

LENGTH: 271

TYPE: PRT

ORGANISM: Oryza sativa

FEATURE:

NAME/KEY: unsure

LOCATION: (1)..(271)

OTHER INFORMATION: unsure at all Xaa locations

FEATURE:

OTHER INFORMATION: Clone ID: PAT\_MRT4530\_14137C.1.pap

US-10-437-963-110035

Query Match 27.5%; Score 690.5; DB 16; Length 271;

Best Local Similarity 66.7%; Pred. No. 3.9e-59;

Matches 134; Conservative 20; Mismatches 28; Indels 19; Gaps 2;

QY 1 HTVDLTIANMMLDSQASDLCKRPE-----ILSMPLPEGKCLLEL 42

DB 71 HSKULTVEANMLDSRAADLDKEERPEFRILQSFSPMGKLLFRVLSPSYKGSVLEL 130

QY 43 GAGIGFT-GLAARACQVIALDFIESAIKKNVINGHYKVKFMCADVTSPILSFPFHS 101

DB 131 GAGIGFTVGTGAKKLAHVIALDFIESAIKKNVINGHYKVKFMCADVTSPILSFPFHS 190

QY 102 LDVIFSNWLLMYLSDVEEVNLVERMLKMLKPGQYIFPESCFHQSDHKKSNPTHREP 161

DB 191 IDLIFSNWLLMYLSDVEEVNLVERMLKMLKPGQYIFPESCFHQSDHKKSNPTHREP 250

QY 162 RYTKAPKECHLDGSGNSYE 182

DB 251 RYTKAPKECHLDGSGNSYE 271

RESULT 12

US-10-424-599-240673

Sequence 240673, Application US/10424599

Publication No. US20040031072A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J

APPLICANT: Kovalic, David K

APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 240673  
LENGTH: 183  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(183)  
OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_59355C.1.pap  
US-10-424-599-240673

Query Match 26.5%; Score 665.5; DB 12; Length 183;  
Best Local Similarity 72.6%; Pred. No. 6.3e-57;  
Matches 122; Conservative 14; Mismatches 25; Indels 7; Gaps 1;

Qy 104 VTFNMLMYSLSDEVENLVERMLKPKGYIFPESCCHQSGDHKRSNPTHYREPRF 163  
Db 20 ITFSNMLMYSLSDEVENLVERMLKPKGYIFPESCCHQSGDHKRSNPTHYREPRF 79

Qy 164 YTKAFKCHLQDGSNGVELSLSCCKTCIGAYVRNKNQISWLMKQVSKDDKQFQPL 223  
Db 80 YTKAFKCHLQDGSNGVELSLSCCKTCIGAYVRNKNQISWLMKQVSKDDKQFQPL 132

Qy 224 DTQYKCNILYRVPFGVYVSTGGYETTKFVSMMLDKPGQKVLVDV 271  
Db 133 DRYEYSKILASERMYGPGVSTGGYETTKFVSMMLDKPGQKVLVDV 180

RESULT 13  
US-10-149-759-50  
Sequence 50, Application US/10149759  
Publication No. US20030157592A1  
GENERAL INFORMATION:  
APPLICANT: Lerchl, Jens  
APPLICANT: Renz, Andreas  
APPLICANT: Ehrhardt, Thomas  
APPLICANT: Reindl, Andreas  
APPLICANT: Clampus, Petra  
TITLE OF INVENTION: Moss Genes from Physcomitrella patens encoding proteins  
TITLE OF INVENTION: Involved in the synthesis of tocopherols and  
FILE REFERENCE: BASE/NAE 1333/99 PCT/US  
CURRENT APPLICATION NUMBER: US/10/149,759  
CURRENT FILING DATE: 2002-10-17  
PRIOR FILING DATE: 2000-12-14  
NUMBER OF SEQ ID NOS: 82  
SOFTWARE: WordPerfect version 6.1  
SEQ ID NO 50  
TYPE: PRT  
ORGANISM: Physcomitrella patens  
US-10-149-759-50

Query Match 26.2%; Score 658; DB 14; Length 168;  
Best Local Similarity 72.9%; Pred. No. 3e-56;  
Matches 121; Conservative 19; Mismatches 26; Indels 0; Gaps 0;

Qy 307 SIGLKCAVEFVADCTKINYPNSPVIYSRDTLHIOKQALFSEYKMLKPGKVLIS 366  
Db 1 SIARKCAVEFVADCTKINYPNSPVIYSRDTLHIOKQALFSEYKMLKPGKVLIS 60

Qy 367 DYCKKAGPPSPFAAYIKQGYDLHDVKEYGQMLKADGVDVLAEDRTQFIRVLKLE 426

Db 61 DYCRAPQTPSAEFAAYIQQGYDLHVSQKYGKMLDEAGFVFEVVAEDRTQFIRVLQRELA 120

Qy 427 TVEKEKDVIFDSFSEYDINDVGMNDKLRRRTAKGEQRMGLFVAKK 472  
Db 121 TTEAGRDQFINDFSEYDINYVSGWKSGLKRCSDNDEQKGLFIAYK 166

RESULT 14  
US-10-369-493-6306  
Sequence 6306, Application US/10369493  
Publication No. US20030233675A1  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
FILE REFERENCE: 38-10(52052)B  
CURRENT APPLICATION NUMBER: US/10/369,493  
CURRENT FILING DATE: 2003-02-28  
PRIOR APPLICATION NUMBER: US 60/360,039  
PRIOR FILING DATE: 2002-02-21  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 6306  
LENGTH: 437  
TYPE: PRT  
ORGANISM: Caenorhabditis elegans  
US-10-369-493-6306

Query Match 20.5%; Score 516.5; DB 15; Length 437;  
Best Local Similarity 33.6%; Pred. No. 1.1e-41;  
Matches 127; Conservative 66; Mismatches 164; Indels 21; Gaps 8;

Qy 104 VIFSNMLMYSLSDEVENLVERMLKPKGYIFPESCCHQSGDHKRSNPTHYREPR 162  
Db 68 LIFNNALSQIITNADLLTDFLNATNATAGCTVIRED-LKDCSDKQVRLTDY---- 122

Qy 163 FYTKAFKCHLQDGSNGVELSLSCCKTCIGAYVRNKNQISWLM---QKVDSDKDKG- 218  
Db 123 ----PDVFTTDSGNNGTGLDLYTDOVER--SNYEQNFLOFIFPRKVPAPTTDAT 175

Qy 219 --PQFLDTSQYKCNILYRVPFGVYVSTGGYETTKFVSMML-DLKPGQKVLVDVCGI 275  
Db 176 ITPRDFLTKQYTNATGIDAYEMFEGVNFISPGGYDENLKIIRFGDFKPGQTMLDIGVGI 235

Qy 276 GGGDFVMASTFVYVVGFDLSVNMISPALESLGK-CAVEFEVADCTKINYPNSPVDVI 334  
Db 236 GGGARQVADFEFGVYVVGFDLSVNMISPALESLGK-CAVEFEVADCTKINYPNSPVDVI 295

Qy 335 YSRDTLHIOKQALFSEYKMLKPGKVLISDYCKKAGPPSPFAAYIKQGYDLHDVVK 394  
Db 296 FSRDCTQHIDPTEKLFSEYKMLKPGKVLISDYCKKAGPPSPFAAYIKQGYDLHDVVK 355

Qy 395 EYQMLKADGVDVLAEDRTQFIRVLKLETVKEKDVIFDSFSEYDINDVGMNDK 454  
Db 356 EIADIANKTGFVNVQNTNTPFRKLEILAEERHLEQNEABEFMSKFTQREDSLSIGWTOK 415

Qy 455 LRRTAKGEQRMGLFVAKK 472  
Db 416 LGYIEKDNENWAFFLAQK 433

RESULT 15  
US-10-602-268-21  
Sequence 21, Application US/10602268  
Publication No. US20040091467A1  
GENERAL INFORMATION:  
APPLICANT: Williams, Deryck J.  
APPLICANT: McIaird, Werty B.  
APPLICANT: Hresko, Michelle Coutu  
APPLICANT: Frevert, Anita M.

```
; APPLICANT: Worthington, Ronald E.
; APPLICANT: Kioek, Andrew P.
; APPLICANT: Davila-Aponte, Jennifer A.
; APPLICANT: Bradley, John D.
; APPLICANT: Xu, Siqun
; TITLE OF INVENTION: NEMATODE PHOSPHOTRANSLAMINE
; FILE REFERENCE: 12557-011001
; CURRENT APPLICATION NUMBER: US/10/602,268
; PRIOR FILING DATE: 2003-06-23
; PRIOR FILING DATE: 2002-06-21
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 437
; TYPB: PRT
; ORGANISM: Caenorhabditis elegans
US-10-602-268-21

Query Match      20.5%; Score 516.5; DB 16; Length 437;
Best Local Similarity 33.6%; Pred. No. 11e-41;
Matches 127; Conservative 66; Mismatches 164; Indels 21; Gaps 8;

Qy 104 VPSNMLMYLSDEE-VENLVERMLKWLKPGGYIPFRESCTHOSGDHKKSNPTHYREPR 162
Db :||: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 68 LIPFNALSQIITNADLLTDFLKNATNATAGGTVIRED-LKDCSDKQVRLTDY---- 122
Db :||: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 163 FYTKAFKECHLDGSGNSYELSLSCKICIGAYVRNKKNQISWLW---QKVDSDKDKG- 218
Db :||: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 123 -----FDVPRITDSGNTGLDLYTDVDEH--SNVVEQNFLDFIPVFRKKVFAPTTDAT 175
Db :||: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 219 --FORFLDTSQYKCNILRYERVPGYVSTGGYETTKFVSM-LDKPCOKVLDVGGGI 275
Db :||: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 176 ITRFDLDTQITNTGIDAYEMFVGFVNPISPGYDENLKIIRFGDFKFGQTMLDIGVI 235
Db :||: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 276 GGGDFYMAETFDVYVVGFDLSVNMISFALERSIGLK-CAVEFEVADCTKINYPDNSFDVI 334
Db :||: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 236 GGGARQVADFQGVHVGIDLSNMLATALERLHEEKDSRVKYSITDALVYQFEDNSFDYV 295
Db :||: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 335 YSRDTLHLODKPALFRSPYKWLPGGKVLISDYCKKAGPPSPFAAYIKORGYDLHDVK 394
Db :||: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 296 PSEDCHQHPDTEKLFPSRIYKALKEGKVLITMGKYGESQDKFKTYVAQRAYFLKX 355
Db :||: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 395 BYGQMLKXAGFDVLAEDRTQPIRVLKELETVEKEKQVPIGDPSEEDYNDIVGGWMDK 454
Db :||: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 356 ELADIANKTGFVNVQTHNMTFRFKXILLERGHLEQNEAEFMSKFTQREDRSLISGWTDK 415
Db :||: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 455 LRETAKGEQRWGLFVAXK 472
Db :||: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 416 LGVIEKDNHNNWFFLAQK 433
Db :||: : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: July 26, 2004, 13:21:07  
Job time : 47 secs



Db 303 SPALHAIGLKCSVEFEVADCTKKEYPDMFTVDVYSRDTILHIQDKPALFRFFYKWLKPG 442  
Qy 361 GKVLISDYCKVAGPPSPFAAYIKRGYDLHDVKEYGQMLKXADAGFVDVLAEDRTQRTV 420  
Db 443 GKVLITDYCRSPKPSDFDAIYIKRGYDLHDVQAYGQMLRAGPFEVIAEDRTQDFMKV 502  
Qy 421 LRKELETVKESKQVFIIDFSEEDYNDIVGWNDDKLRRTAKGRWGLFVAKK 472  
Db 503 LKRELDVAKESKEFISDFSKEDYEDIIIGWKSILARSSGQKWLGTAKR 554

RESULT 2  
F96525  
protein T1N15.23 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: F96525  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizcar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.K.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: F96525  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-374 <STO>  
A:Cross-references: GB:AR005173; NID:98778697; PIDN:AAF97905.1; GSPDB:GN00141  
C:Genetics:  
A:Gene: T1N15.23  
A:Map position: 1

Query Match 51.5%; Score 1294.5; DB 2; Length 374;  
Best Local Similarity 66.8%; Pred. No. 3e-89;  
Matches 245; Conservative 30; Mismatches 33; Indels 59; Gaps 2;  
Qy 166 KAFKECHLDGSGNSYELLSCKICIGAYVRNKKQNIQISMLQKVDSDKDGQFQRLDT 225  
Db 8 QVQECQCDASGNSFELSMGCKICIGAYVRNKKQNIQIMQKVSVDNDKQFQRLDN 67  
Qy 226 SQKNSILYERVPFGVSTGCGY-----ETTKK 255  
Db 68 VQKSGILYERVPFGVSTGCGFNSILLSSYGYTYCLSVIPMLPSLTETKE 127  
Qy 256 FVSMDLKPGQKVLVDGCGIGGDFYMAETFDVVGFDLSVNMISFALERSIGLKCAVE 315  
Db 128 FVAMDLKPGQKVLVDGCGIGGDFYMAENFDVHVVGIDLSVNMISFALERAIGKCSVE 187  
Qy 316 FEVADCTKINYPNSFDVYSRDTILHIQDKPALFRFPKMLKPGKVLISDYCKKAGPP 375  
Db 188 FEVADCTKINYPNSFDVYSRDTILHIQDKPALFRFPKMLKPGKVLITDYCRSAETP 247  
Qy 376 SPFAAYIKRGYDLHDVKEYGQMLKXADAGFVDVLAEDRTQRTVLRKELETVKESKQV 435  
Db 248 SPFAEYIKRGYDLHDVQAYGQMLKXADAGFVDVIAEDRTQFVQLRRELEKSKESKEEF 307  
Qy 436 ISDFSE-----EDYNDIVGWNDDKLRRTAKGRWGLFVAKK 466  
Db 308 ISDFSEVKTTLTNSFQNTCLLCHDHKLQQLQDYNDIVGGSAXLERTASGEQKMG 367  
Qy 467 LFVAKK 473  
Db 368 LFADKK 374

RESULT 3  
T29330  
hypothetical protein F54D11.1 - Caenorhabditis elegans  
Query Match 16.4%; Score 413; DB 2; Length 495;  
Best Local Similarity 28.6%; Pred. No. 4.3e-23;

C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T29330  
R:Pauley, A.; Gattung, S.  
submitted to the EMBL Data Library, July 1996  
A:Description: The sequence of C. elegans cosmid F54D11.  
A:Reference number: 220606  
A:Accession: T29330  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-437 <PAU>  
A:Cross-references: EMBL:U64834; PIDN:AA04824.1; GSPDB:GN00023; CESP:F54D11.1  
A:Experimental source: strain Bristol N2; clone F54D11  
C:Genetics:  
A:Gene: CESP:F54D11.1  
A:Map position: 5  
A:Introns: 36/2; 76/3; 108/1; 315/3; 398/3  
Query Match 20.5%; Score 516.5; DB 2; Length 437;  
Best Local Similarity 33.6%; Pred. No. 6.5e-31;  
Matches 127; Conservative 66; Mismatches 164; Indels 21; Gaps 8;  
Qy 104 VIFSNMLMLYLSDES-VENUVERMLKWKPGGHIFFRESCFHSGDHKKSHPHYRER 162  
Db 68 LIFNNALSQIITNADLLTDFLKNATNATAGGTVIRED-LKDCSKRQVARLTDI---- 122  
Qy 163 FYTKAPKECHLDGSGNSYELLSCKICIGAYVRNKKQNIQISMLW---QKVDSDKDKG- 218  
Db 123 -----FDVPRITDSDGNTGLDLTVDOVEH--SNYVEQNFLDFIVFRKCVFAPTDDAT 175  
Qy 219 --PQRLDTSQYKCHSILRYERVPFGVSTGCGYVETTKFVSM-LDKPGQKVLVDGCGI 275  
Db 176 ITRDFLDKQTVNTGIDAYEMFGVWFISPGYDENLKIIFGDFKFGQTMLDIGVGI 235  
Qy 276 GGGDFYMAETFDVVGFDLSVNMISFALERSIGLK-CAVEPEVADCTKINYPNSFDV 334  
Db 236 GGGARQVADRFVHVHVGIDLSNMLALALERLHBEKDSRVKYSITDALVTQFEDNSFDIV 295  
Qy 335 YSRDTILHIQDKPALFRFPKMLKPGKVLISDYCKKAGPPSPFAAYIKRGYDLHDV 394  
Db 296 FSRDCHQHPDYKLFPSRIYKALKPGKVLITWYKGYGQSKFKTYVAGRAYFLANLK 355  
Qy 395 EYQMLKXADAGFVDVLAEDRTQRTVLRKELETVKESKQVFIIDFSEEDYNDIVGWN 454  
Db 356 EIADIANKTQFVNVQFNTNTPRPFKEILLERHLSQNEAEFMSKFTQERDLSISGWTDK 415  
Qy 455 LRRTAKGEQKWLFLVAKK 472  
Db 416 LGVIEKNHNNWFFLAQK 433

RESULT 4  
T27936  
hypothetical protein ZK622.3 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 04-Mar-2000  
C:Accession: T27936  
R:Leimbach, D.  
submitted to the EMBL Data Library, November 1995  
A:Description: The sequence of C. elegans cosmid ZK622.  
A:Reference number: 220443  
A:Accession: T27936  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-495 <LEI>  
A:Cross-references: EMBL:U99998; PIDN:AAA81102.1; CESP:ZK622.3  
C:Genetics:  
A:Gene: CESP:ZK622.3  
A:Introns: 34/2; 97/2; 222/3; 310/3  
C:Superfamily: Caenorhabditis elegans hypothetical protein ZK622.3  
Query Match 16.4%; Score 413; DB 2; Length 495;  
Best Local Similarity 28.6%; Pred. No. 4.3e-23;



Db 49 WEDVGHGHHGYYGTYRDRQAQIDLIKELLAWAVPQNSAKP-RKILDLGGGIGG 107  
 QY 278 GDFMAETFDVEVGGDLSVNMISPALE--RSIGLKCAVEFVADCTKINYPNSPDVY 335  
 Db 108 SSLYLAAQQAQVAVGASLSVPQVRAGRRARALGIGSTCOFQVANALDLPASDFDWW 167  
 QY 336 SRDILTIHQKPALEFRFYKWLKPGGKVLISDYCKA-----GPPSPFAAYIKORGYDL 390  
 Db 168 SLESGEHMKNAQPLQEAARVLPKPGGRLILATWCHRPIDPGNGPLTADERRHL-QAIYDV 226  
 QY 391 H-----DVKEYGOMLADGPDVLAED 412  
 Db 227 YCLPVVSLPYEAIAECGFGHKTAD 254

RESULT 8  
 T01572  
 sterol 24-C-methyltransferase (EC 2.1.1.41) - maize  
 N;Alternate names: (S)-adenosyl-L-methionine:delta 24-sterol methyltransferase  
 C;Species: Zea mays (maize)  
 C;Date: 19-Feb-1999 #sequence\_revision 19-Feb-1999 #text\_change 03-Jun-2002  
 C;Accession: T01572  
 R;Tong, X.; Nes, W.D.  
 submitted to the EMBL Data Library, January 1998  
 A;Reference number: Z14350  
 A;Accession: T01572  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-344 <TON>  
 A;Cross-references: EMBL:AF045570; NID:G2909845; PIDN:AAC04265.1; PID:G2909846  
 A;Experimental source: strain B73  
 C;Genetics:  
 A;Gene: SMT  
 C;Superfamily: 24-sterol C-methyltransferase; bioC homology  
 C;Keywords: methyltransferase; S-adenosylmethionine  
 F:101-205/Domain: bioC homology <BIOC>

Query Match 7.9%; Score 199; DB 2; Length 344;  
 Best Local Similarity 34.5%; Pred. No. 2.8e-07;  
 Matches 60; Conservative 26; Mismatches 72; Indels 16; Gaps 7;  
 QY 254 KEFVSM-LDLKPGQKVLVGGGIGGDFYMAETFDVVGFDLSVNMISPA--LERSIGL 310  
 Db 89 EHFLALQLGLKPGKVLVGGGIGGGLREIARFSTSVTGLNNNEYQITRGKELNRLAGI 148  
 QY 311 KCAVEFVADCTKINYPNSPDVYSRDILTIHQKPALEFRFYKWLKPGGKVLISDYCK 370  
 Db 149 SGTCDVFVKADEFMPPDDNTFDVYATCATCHAPDPVGCYKEIYRVLPKPGQCFVYEW- 207  
 QY 371 KAGPPSPFAAY--IK---ORGYDLHDVKEYGOML---KDAGFVDV----LAED 412  
 Db 208 VTDHYDPNNAHTRIKDBIELGNGLPDIRSTRQCLRAVKDAGFEVVDKDLAED 261

RESULT 9  
 T04138  
 sterol 24-C-methyltransferase (EC 2.1.1.41) ESMRL, endosperm - maize  
 N;Alternate names: C-24 sterol methyltransferase  
 C;Species: Zea mays (maize)  
 C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 03-Jun-2002  
 C;Accession: T04138  
 R;Greenok, R.J.; Galbraith, D.W.; Dellapenna, D.  
 Plant Mol. Biol. 34: 891-896, 1997  
 A;Title: Characterization of zea-mays endosperm C-24 sterol methyltransferase - one of 2  
 A;Reference number: Z09668; MUID:97435974; PMID:9290641  
 A;Accession: T04138  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-344 <GRE>  
 A;Cross-references: EMBL:U79669; NID:G1899059; PIDN:AAB70886.1; PID:G1899060  
 A;Experimental source: endosperm  
 C;Genetics:

A;Gene: ESMT1  
 C;Function:  
 A;Description: methyltransferase  
 C;Superfamily: 24-sterol C-methyltransferase; bioC homology  
 C;Keywords: methyltransferase; S-adenosylmethionine  
 F:101-205/Domain: bioC homology <BIOC>  
 Query Match 7.9%; Score 198; DB 2; Length 344;  
 Best Local Similarity 34.4%; Pred. No. 3.3e-07;  
 Matches 59; Conservative 26; Mismatches 74; Indels 14; Gaps 6;  
 QY 254 KEFVSM-LDLKPGQKVLVGGGIGGDFYMAETFDVVGFDLSVNMISPA--LERSIGL 310  
 Db 89 EHFLALQLGLKPGKVLVGGGIGGGLREIARFSTSVTGLNNNEYQITRGKELNRLAGI 148  
 QY 311 KCAVEFVADCTKINYPNSPDVYSRDILTIHQKPALEFRFYKWLKPGGKVLISDYC- 369  
 Db 149 SGTCDVFVKADEFMPPDDNTFDVYATCATCHAPDPVGCYKEIYRVLPKPGQCFVYEW- 208  
 QY 370 KAGPPSPFAAYIK---ORGYDLHDVKEYGOML---KDAGFVDV----LAED 412  
 Db 209 TDDHYDPNNAHTRIKDBIELGNGLPDIRSTRQCLRAVKDAGFEVVDKDLAED 261

RESULT 10  
 T06780  
 probable sterol 24-C-methyltransferase (EC 2.1.1.41) - soybean  
 N;Alternate names: S-adenosyl-L-methionine:delta24-sterol-C-methyltransferase  
 C;Species: Glycine max (soybean)  
 C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 03-Jun-2002  
 C;Accession: T06780  
 R;Shi, J.; Gonzales, R.A.; Bhattacharyya, M.K.  
 J. Biol. Chem. 271, 9384-9389, 1996  
 A;Title: Identification and characterization of an S-adenosyl-L-methionine: delta 24-ster  
 A;Reference number: Z15807; MUID:96199190; PMID:8621604  
 A;Accession: T06780  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-367 <SHI>  
 A;Cross-references: EMBL:U43683; NID:G1399379; PIDN:AAB04057.1; PID:G1399380  
 A;Experimental source: cultivar Williams 82; etiolated hypocotyls  
 C;Superfamily: 24-sterol C-methyltransferase; bioC homology  
 C;Keywords: methyltransferase; S-adenosylmethionine  
 F:122-226/Domain: bioC homology <BIOC>

Query Match 7.7%; Score 193; DB 2; Length 367;  
 Best Local Similarity 27.2%; Pred. No. 8.5e-07;  
 Matches 63; Conservative 35; Mismatches 90; Indels 44; Gaps 7;  
 QY 260 LDLKPGQKVLVGGGIGGDFYMAETFDVVGFDLSVNMISPA--LERSIGLKCAVEFE 317  
 Db 117 LGKPGQKVLVGGGIGGGLREIARFSTSVTGLNNNEYQITRGKELNRLAGVDTCTNEV 176  
 QY 318 VADCTKINYPNSPDVYSRDILTIHQKPALEFRFYKWLKPGGKVLISDYC--KAGPP 375  
 Db 177 KADFMKMPDPSFDVYATCATCHAPDPVGCYKEIYRVLPKPGQYFAAYEWCMTDSFDFO 236  
 QY 376 SPEFAAYIK---ORGYDLHDVVK---EYGOMLKADGFEVVDVLAEDR----- 413  
 Db 237 NPEHOK-IAKEIEIGDGLPDILRTAKCLEALKQAGFEVWEKDLAVDPSLPWYLPDLKSH 295  
 QY 414 -----TEQIRVLRKELETVEKEKDVISDPSREDYNDIVGG 450  
 Db 296 FSLSSPFLTAVGRLFTTKNNVKV--EYVGLAPKGLSLVQDFLEKAGRLVEG 345

RESULT 11  
 AC2071  
 hypothetical protein all2121 (imported) - Nostoc sp. (strain PCC 7120)  
 C;Species: Nostoc sp. PCC 7120  
 A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp.  
 C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
 C;Accession: AC2071





Db 200 AKERVKELGLEDKIEIQLQYQDLRF-ENYFDKVVSVCMFPHVGVKGNLGLVFMVKQVLK 258  
QY 359 PGGKVLII-SDYCKKAGPPSEFAAYIKQGYDLHDKVYGOHMKDAGFVDVLAEDRTEOF 417  
Db 259 PGGKVLII-SILAMPEGKTNADIKYIPGGY-LPSLRVVSAMSEWDPHLLASLRMHY 317  
QY 418 IRVLKSLVETVEKOVFISDFSESDYNDIVGMDNKLRTAKGE--QRWGLFV 469  
Db 318 AKTL-----DLMDENFNKVL-----DKVREKYDEEFIRMOLYL 351

RESULT 15

S18533  
eryG protein - Saccharopolyspora erythraea  
C;Species: Saccharopolyspora erythraea  
C;Date: 07-Apr-1994 #sequence\_revision 07-Apr-1994 #text\_change 22-Oct-1999  
C;Accession: S18533; S16747  
R;Haydock, S.F.; Dowson, J.A.; Dhillon, N.; Roberts, G.A.; Cortes, J.; Leadlay, P.P.  
Mol. Gen. Genet. 230, 120-128, 1991  
A;Title: Cloning and sequence analysis of genes involved in erythromycin biosynthesis in  
methylintrasferases.  
A;Reference number: S18530; MUID:92079886; PMID:1840640  
A;Accession: S18533  
A;Molecule type: DNA  
A;Residues: 1-306 <HAY>  
A;Cross-references: EMBL:X60379; NID:948941; PIDN:CA42929.1; PID:G581650  
C;Genetics:  
A;Gene: eryG  
A;Start codon: GTG  
C;Superfamily: bioC homology  
F;82-187/Domain: bioC homology <BIOC>

Query Match 7.2%; Score 180.5; DB 2; Length 306;  
Best Local Similarity 24.5%; Pred. No. 5.8e-06;  
Matches 64; Conservative 45; Mismatches 95; Indels 57; Gaps 11;  
QY 206 W-LWQKV--DSKDDKGFQRPDLTQYKCNSTLYRVERVFGPGYVSTGGY-----ETTK 254  
Db 18 WHVWTRVPSRRALAYALFADD-----HEATTEGAYINLGTWPGCNGLEHNG 67  
QY 255 EFVSMI-----DLKPGQKVLGVCGIGGDFYMAET-FDVEVVGFDLSVNMISPALEERS-- 307  
Db 68 ELANQLAERAGISGEDEVLDVGFUGAQDFPWLSTRKPARIVGVDLTPSHVRIASERAER 127  
QY 308 IGLKCAVEEVADCTKINYPDNPDIYSDTILHIOKPALEFRSFYKMLKPGKVLISD 367  
Db 128 ENVQDRLLQFAGSATDLPFGAETFDRTVTSLESALRYEPTDFFKGAFEVLKPGVLAIGD 187  
QY 368 YCK-----KAGPPSPPEAAVIKQGYDLH-----DVKEYGQKLDAGFVDV--- 408  
Db 188 IIPDLRERPGSDPPK-----LAPQSGSLSGGIPVENWVPRETYAKQLEBAGFVDVVK 242  
QY 409 ----LAEDRTEQFINVLKKE 424  
Db 243 SVRDNWPWLDVYLKQLQDE 263

Search completed: July 26, 2004, 13:15:54  
Job time : 19 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2004, 13:11:16 ; Search time 13 Seconds  
(without alignments)  
1894.553 Million cell updates/sec

Title: US-10-031-331b-40

Perfect score: 2315

Sequence: 1 HTVDLTFEMLDQASDL.....KURRTAKGQRWGLFVAKK 473

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2204	87.6	494	1	PEAM SPIOL
2	2077	82.6	475	1	PMW2 ARATH
3	2041	81.2	490	1	PMW3 ARATH
4	2011	80.0	491	1	PMW4 ARATH
5	176.5	7.0	251	1	UBIE SHON
6	173.5	6.9	348	1	GROM ARATH
7	167.5	6.7	241	1	UBIE STAM
8	161.5	6.4	260	1	UBIE VIRCH
9	161.5	6.4	381	1	CPA ECOLI
10	160.5	6.4	251	1	UBIE ECOLI
11	160.5	6.4	259	1	UBIE VIBPA
12	159.5	6.3	378	1	ERG6 SCHPO
13	158.5	6.3	251	1	UBIE ECOLI
14	158.5	6.3	260	1	UBIE VIBVU
15	157	6.2	245	1	UBIE NEIMA
16	156.5	6.2	241	1	UBIE STAP
17	156	6.2	245	1	UBIE NEIMB
18	155.5	6.2	237	1	UBIE LISMO
19	155.5	6.2	238	1	UBIE OCEIH
20	154.5	6.1	244	1	YXBB BACSU
21	154.5	6.1	256	1	UBIE PBPBK
22	154.5	6.1	376	1	ERG6 CANAL
23	153.5	6.1	253	1	UBIE XANCP
24	153	6.1	382	1	ERG6 YEAST
25	151.5	6.0	251	1	UBIE SALTY
26	151.5	6.0	251	1	UBIE YERPE
27	149.5	5.9	237	1	UBIE PASMA
28	149.5	5.9	268	1	UBIE PASMU
29	148.5	5.9	253	1	UBIE XANAC
30	148.5	5.9	256	1	UBIE PSESM
31	147.5	5.9	237	1	UBIE BACCR
32	147.5	5.9	256	1	UBIE PSEPT
33	146.5	5.8	250	1	UBIE CONSU

## RESULT 1

ID	PEAM SPIOL	STANDARD;	PRT;	494 AA.
AC	Q9M571;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Phosphoethanolamine N-methyltransferase [EC 2.1.1.103].			
GN	PEAMT.			
OS	Spinacia oleracea (Spinach).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
OC	Caryophyllales; Amaranthaceae; Spinacia.			
OX	NCBI_TaxID=3562;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND CHARACTERIZATION.			
RC	STRAIN=sv. Savoy hybrid 612; TISSUE=Leaf;			
RX	MEDLINE=20261526; PubMed=10799484;			
RA	Nuccio M.L., Ziemak M.J., Henry S.A., Weretinsky E.A., Hanson A.D.;			
RT	"Phosphoethanolamine N-methyltransferase from spinach: cDNA cloning by			
RT	complementation in <i>Schizosaccharomyces pombe</i> and characterization of			
RT	the recombinant enzyme."			
RL	J. Biol. Chem. 275:14095-14101(2000).			
CC	-!- FUNCTION: Catalyzes N-methylation of phosphoethanolamine,			
CC	phosphomonomethyl ethanolamine and phosphodimethyl ethanolamine, the			
CC	three methylation steps required to convert phosphoethanolamine to			
CC	phosphocholine. Mediates a key step in the biosynthesis of			
CC	choline, a precursor of the osmoprotectant glycine betaine. Has an			
CC	optimal pH of 7.8 to 8.5. Has no ethanolamine- or			
CC	phosphatidylethanolamine-N-methyltransferase activity.			
CC	-!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + ethanolamine			
CC	phosphate = S-adenosyl-L-homocysteine + N-methylethanolamine			
CC	-!- ENZYME REGULATION: Inhibited by phosphocholine but not by choline,			
CC	glycine betaine, monomethylethanolamine or dimethylethanolamine.			
CC	-!- SUBUNIT: Monomer.			
CC	-!- SUBCELLULAR LOCATION: Cytoplasmic.			
CC	-!- INDUCTION: By salt stress.			
CC	-!- SIMILARITY: Belongs to the methyltransferase superfamily.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	EMBL; AF237633; AAF61950.1; --			
DR	InterPro; IPR001601; Methyltransf.			
DR	InterPro; IPR000051; SAM bind.			
KW	Methyltransferase; transferase; Repeat.			
FT	DOMAIN 1..53.....I62..... SAM-BINDING 1..			
FT	DOMAIN 282..389 SAM-BINDING 2.			
FT	SEQUENCE 494 AA; 56361 MW; 7F2537C8E4B8413B CRC64;			

## ALIGNMENTS

34	146	5.8	231	1	UBIE STRCO
35	145.5	5.8	259	1	UBIE HAEDU
36	143.5	5.7	224	1	UBIE BACST
37	143.5	5.7	256	1	UBIE PSEAZ
38	141	5.6	345	1	YI37_STRFR
39	140.5	5.6	286	1	MMAL_MYCTU
40	140	5.6	244	1	UBIE NITEU
41	140	5.6	246	1	UBIE MICLU
42	138	5.5	239	1	UBIG XANAC
43	136.5	5.4	230	1	UBIE CORBF
44	135.5	5.4	252	1	UBIE LACLA
45	134	5.3	231	1	UBIE STRAW

Q9xap8	streptomyces
P59911	haemophilus
O86169	bacillus st
O98uc0	pseudomonas
P20187	streptomyces
P94922	mycobacteri
O81zz8	nitrosomona
O66128	micrococci
Q8pk00	xanthomonas
Q8feb3	corynebacte
P49016	lactococcus
Q81zx2	streptomyces

```

Query Match      87.6%; Score 2204; DB 1; Length 494;
Best Local Similarity 86.4%; Pred. No. 8.9e-159;
Matches 408; Conservative 34; Mismatches 30; Indels 0; Gaps 0;

QY 1 HTVLTITVAMMLDSQASDLKBERPHTLSMLPLEGKCLLELGAGIGFTGELAAGQV 60
DB 22 HSDVLTVEAMMLDSQASDLKBERPHTLSMLPLEGKCLLELGAGIGFTGELAAGQV 81
QY 61 IALDFIESAIGKNEVINGHYKVKFPCADVTSTLSPFPSPHSDVIFSNWLLMYLSDEVE 120
DB 82 IALDFIESAIGKNEVINGHYKVKFPCADVTSTLSPFPSPHSDVIFSNWLLMYLSDEVE 141
QY 121 NLVERMLKWLKPGGYIFFRSCPHQSDHKKSNPHTYREPRFTYKAPKCECHLQDGSNS 180
DB 142 RLVERMLKWLKPGGYIFFRSCPHQSDHKKSNPHTYREPRFTYKAPKCECHLQDGSNS 201
QY 181 YELSLSCCKTGAVYKVKNNQNOISLWKQVDSKDKGFORPLDTSOYKCNSTLYRYVFP 240
DB 202 YELSLSCCKTGAVYKVKNNQNOISLWKQVDSKDKGFORPLDTSOYKCNSTLYRYVFP 261
QY 241 GPGYVSTGGYETTFKVFMSMLDLKPGQKVLVGGCGIGGDFYMAETDVEVGFDSVNM 300
DB 262 GPGYVSTGGYETTFKVFMSMLDLKPGQKVLVGGCGIGGDFYMAETDVEVGFDSVNM 321
QY 301 SPALERSIGLCAVEFEVADCTKINYPDSFDVYISRDITLHQDKPALFRFYKWLKPG 360
DB 322 SPALERSIGLCAVEFEVADCTKINYPDSFDVYISRDITLHQDKPALFRFYKWLKPG 381
QY 361 GKVLSDYCKAGPSPSPFAAYIKORGYDLHDVKEVGMKQKAGFVDVLAERTQPIRV 420
DB 382 GKVLSDYCKAGPSPSPFAAYIKORGYDLHDVKEVGMKQKAGFVDVLAERTQPIRV 441
QY 421 LRKELETKEKGVPIIDFSEEDYNDIVGGNNDKLARTKGEORGLFVAKK 472
DB 442 LQKELDALEKQKDFIDFSEEDYNDIVGGNNDKLARTKGEORGLFVAKK 493

RESULT 2
PBM2_ARATH STANDARD; PRT: 475 AA.
ID PBM2_ARATH STANDARD; PRT: 475 AA.
AC Q94H0; Q9LP63; Q9LP64;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 15-MAR-2004 (Rel. 43, Last annotation update)
DE Putative phosphoethanolamine N-methyltransferase 2 (BC 2.1.1.103).
GN NMT2 OR AT1G48600 OR T1N15.20 OR T1N15.22/T1N15.23.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia; PubMed=11130712;
RX MEDLINE=21016719;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altarfi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Egu P., Feldlyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marzilli A.,
RA Millican J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tabunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vayenberg M., Ysotokajia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT *Sequence and analysis of chromosome 1 of the plant Arabidopsis
thaliana.*;

```

```

RL Nature 408:816-820(2000).
RP SEQUENCE FROM N.A.
RX STRAIN=cv. Columbia; PubMed=14593172;
RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,
RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
RA Karlin-Newmann G., Liu S.X., Lam B., Sekano H., Wu T., Yu G.,
RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
RA Chan M.M., Tang C.C., Onodera C.S., Deng J.H., Akiyama K., Ansari Y.,
RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,
RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.P.,
RA Hayashizaki Y., Johnson-Hopson C., Huan V.W., Iida K., Karnes M.,
RA Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
RA Satou M., Tenen R., Vayenberg M., Wallender E.K., Wong C., Yamamura Y.,
RA Yuan S., Shinozaki K., Davis R.M., Theologis A., Ecker J.R.;
RT *Empirical analysis of transcriptional activity in the Arabidopsis
genome.*;
RL Science 302:842-846(2003).
CC -1- FUNCTION: Catalyzes N-methylation of phosphoethanolamine,
CC phosphomonomethylethanolamine and phosphodimethylethanolamine, the
CC three methylation steps required to convert phosphoethanolamine to
CC phosphocholine (By similarity).
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + ethanolamine
CC phosphate = S-adenosyl-L-homocysteine + N-methylethanolamine
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the methyltransferase superfamily.
CC -1- CAUTION: Ref1 sequences differ from that shown due to erroneous
CC gene model prediction.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; AC020889; AAF79704.1; ALT SEQ.
CC EMBL; AC020889; AAF79705.1; ALT SEQ.
CC EMBL; AF428454; AAL16223.1; -.
CC EMBL; AY063866; AAL36222.1; -.
CC InterPro; IPR001601; Methyltransf.
CC InterPro; IPR000051; SAM bind.
CC InterPro; IPR004033; UbiE/COO5 Metrf.
CC Pfam; PF01209; Ubie_methyltran; 1.
CC Transferrase; Methyltransferase; Repeat.
CC DOMAIN 34 143 SAM-BINDING 1.
CC FT DOMAIN 263 370 SAM-BINDING 2.
CC SEQUENCE 475 AA; 54018 MW; 04070544ADE80DA8 CRC64;

Query Match      82.6%; Score 2077; DB 1; Length 475;
Best Local Similarity 80.1%; Pred. No. 3.2e-149;
Matches 379; Conservative 47; Mismatches 47; Indels 0; Gaps 0;

QY 1 HTVLTITVAMMLDSQASDLKBERPHTLSMLPLEGKCLLELGAGIGFTGELAAGQV 60
DB 3 HSDVLTVEAMMLDSQASDLKBERPHTLSMLPLEGKCLLELGAGIGFTGELAAGQV 62
QY 61 IALDFIESAIGKNEVINGHYKVKFPCADVTSTLSPFPSPHSDVIFSNWLLMYLSDEVE 120
DB 63 IALDFIESAIGKNEVINGHYKVKFPCADVTSTLSPFPSPHSDVIFSNWLLMYLSDEVE 122
QY 121 NLVERMLKWLKPGGYIFFRSCPHQSDHKKSNPHTYREPRFTYKAPKCECHLQDGSNS 180
DB 123 LNAERMIGWKPFGGYIFFRSCPHQSDHKKSNPHTYREPRFTYKAPKCECHLQDGSNS 182
QY 181 YELSLSCCKTGAVYKVKNNQNOISLWKQVDSKDKGFORPLDTSOYKCNSTLYRYVFP 240
DB 183 FELSHWGWCKICGAVYKVKNNQNOISLWKQVDSKDKGFORPLDTSOYKCNSTLYRYVFP 242

```

CC 241 GPGVSTGCTGTTTKEFVSMGLDKPGQKVLVDVGGGDFMAETFDVVGFDLSVNNI 300  
 CC 243 GEGVSTGCTGTTTKEFVSMGLDKPGQKVLVDVGGGDFMAETFDVVGFDLSVNNI 302  
 CC 301 SPALERSIGLCAVSEFVADCTKINYPNSDVTYSRDTLHODKALFRSPYKWLKPG 360  
 CC 303 SPALERAIGLCAVSEFVADCTKINYPNSDVTYSRDTLHODKALFRSPYKWLKPG 362  
 CC 361 GKVLISDYCKGAGPSPFAAYIKRGYDLHVKYQMLKADGAFVDVLAORTQPIRV 420  
 CC 363 GKVLITDYCRSAETSPFAAYIKRGYDLHVKYQMLKADGAFVDVLAORTQPIRV 422  
 CC 421 LRKELETVEKEDVTSFSDSEYNDIVGGWNDKLRRTAKGEORGLFVAKKK 473  
 CC 423 LRRELEKVEKEEFSDSEYNDIVGGWNDKLRRTAKGEORGLFVAKKK 475

## RESULT 3

PEM1 ARATH STANDARD; PRT; 490 AA.  
 AC Q9C6B9; Q9C9V1;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Putative phosphoethanolamine N-methyltransferase 3 (EC 2.1.1.103).  
 GN NW3 OR AT1G73600 OR P605.1 OR P25P22.1.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsi.  
 NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=21016719; PubMed=11130712;  
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,  
 RA White O., Alonso J., Altafi H., Araujo R., Bowman C.B., Brooks S.Y.,  
 RA Buehler S., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,  
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,  
 RA Dunn P., Feldblyum T.V., Feng J.-D., Feng B., Fujii C.Y.,  
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,  
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin B.,  
 RA Kim C.J., Koo H.L., Krenetskaia I., Kurtz D.B., Kwan A., Lam B.,  
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,  
 RA Millican J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,  
 RA Paj G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
 RA Sakao H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
 RA Utterback T., Van Aken S., Vaysberg M., Vyeotskaia V.S., Walker M.,  
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
 RA "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
 thaliana.";  
 RL Nature 408:816-820(2000).  
 RN [2]  
 RP CONCEPTUAL TRANSLATION.  
 RA Schneider M.;  
 CC Unpublished observations (MAY-2002).  
 CC -1- FUNCTION: Catalyzes N-methylation of phosphoethanolamine,  
 CC three methylation steps required to convert phosphoethanolamine to  
 CC phosphocholine (By similarity).  
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + ethanolamine  
 CC phosphate = S-adenosyl-L-homocysteine + N-methyl ethanolamine  
 CC phosphate.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: Belongs to the methyltransferase superfamily.  
 CC -1- CAUTION: Ref.1 (AGS1806) sequence differs from that shown due to  
 CC erroneous gene model prediction.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use.

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).

CC -----  
 CC EMBL; AC079676; AAGS1806.1; ALT\_SEQ.  
 CC EMBL; AC012679; AAGS2075.1; -; -;  
 CC InterPro; IPR001601; Methyltransf.  
 CC InterPro; IPR000051; SAM bind.  
 CC K0 Transferrase; Methyltransferase; Repeat.  
 CC FT DOMAIN 49 158 SAM-BINDING 2.  
 CC FT DOMAIN 278 385 SAM-BINDING 2.  
 CC SQ SEQUENCE 490 AA; 56368 MW; 77DFAF8C9C41CB CRC64;

Query Match 81.2%; Score 2041; DB 1; Length 490;  
 Best Local Similarity 78.4%; Pred. No. 1.7e-146;  
 Matches 370; Conservative 53; Mismatches 49; Indels 0; Gaps 0;

QY 1 HTVDLTIANWMLDSQASDLDEXEPRLTSMPLPELEGKCLLELGAIGRTTGLAKAGQV 60  
 DB 18 HSVGLSVEAMWMLDSKASDLDEXEPRLTSMPLPELEGKCLLELGAIGRTTGLAKAGQV 77  
 QY 61 IALDPTESAIKKNVINGHYKVKFMCAVTSPTLSFPPSHLDVIFSNWLLMYLSDQEV 120  
 DB 78 IAVDTIESVIKKNENINGHYKVKFLCADVTSNNMFPNESMDLIFSNWLLMYLSDQEV 137  
 QY 121 NIVEMLWKLKPGYIFPSCFHSQSDHKKSNPTHYRPRFTYKAFKECHLQDSGNS 180  
 DB 138 DLAKMLOMTQVGVYIFPSCFHSQSDHKKSNPTHYRPRFTYKAFKECHLQDSGNS 197  
 QY 181 YELSLSCCKCTGAYVRNKNQNSIWLQKVDKSGDKGFORFLDTSTQYKNSILYERVF 240  
 DB 198 YELSLVSCCKCTGAYVRNKNQNSIWLQKVDKSGDKGFORFLDTSTQYKNSILYERVF 257  
 QY 241 GCGYSTGCTGTTTKEFVSMGLDKPGQKVLVDVGGGDFMAETFDVVGFDLSVNNI 300  
 DB 258 GCGYSTGCTGTTTKEFVSMGLDKPGQKVLVDVGGGDFMAETFDVVGFDLSVNNI 317  
 QY 301 SPALERSIGLCAVSEFVADCTKINYPNSDVTYSRDTLHODKALFRSPYKWLKPG 360  
 DB 318 SPALERAIGLCAVSEFVADCTKINYPNSDVTYSRDTLHODKALFRSPYKWLKPG 377  
 QY 361 GKVLISDYCKGAGPSPFAAYIKRGYDLHVKYQMLKADGAFVDVLAORTQPIRV 420  
 DB 378 GKVLITDYCRSAETSPFAAYIKRGYDLHVKYQMLKADGAFVDVLAORTQPIRV 437  
 QY 421 LRKELETVEKEDVTSFSDSEYNDIVGGWNDKLRRTAKGEORGLFVAKK 472  
 DB 438 LRRELEKVEKEEFSDSEYNDIVGGWNDKLRRTAKGEORGLFVAKK 489

## RESULT 4

PEM1 ARATH STANDARD; PRT; 491 AA.  
 ID PEM1 ARATH STANDARD; PRT; 491 AA.  
 AC Q9FR44; Q9LVH3;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Phosphoethanolamine N-methyltransferase 1 (EC 2.1.1.103) (PEAMT 1)  
 DE (AtNMT1).  
 GN NMT1 OR AT3G17990/AT3G18000 OR MB5.19 OR MB5.21/MB5.22.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.  
 NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Landsberg erecta; TISSUE=Seedling;  
 RX MEDLINE=20567827; PubMed=11115895;  
 RA Bolognese C.P., McGraw P.;  
 RA "The isolation and characterization in yeast of a gene for Arabidopsis  
 S-adenosylmethionine:phospho-ethanolamine N-methyltransferase.";





CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AF015852; AAN57171.1; -  
 DR TIGR; S04199; -  
 DR HAVAP; MP\_01813; -; 1.  
 DR InterPro; IPR001601; Methyltransf.  
 DR InterPro; IPR000051; SAM bind.  
 DR InterPro; IPR004034; UniProt/MetTransf.  
 DR InterPro; IPR004033; UniProt/COG5\_MetTrf.  
 DR Pfam; PF01209; UbiE\_methyltransf; 1.  
 DR PROSITE; PS01183; UBI1\_1; 1.  
 DR PROSITE; PS01184; UBI1\_2; 1.  
 KW Menquinone biosynthesis; Ubiquinone biosynthesis; Transferase;  
 KW Methyltransferase; Complete proteome.  
 SQ SEQUENCE 251 AA; 28065 MW; EP0AFC3CF06B2141 CRC64;

Query Match 7.0%; Score 176.5; DB 1; Length 251;  
 Best Local Similarity 28.3%; Pred. No. 3.5e-06;  
 Matches 63; Conservative 43; Mismatches 76; Indels 41; Gaps 12;

OY 245 VSTGQYETKTF-VSMLDLPGKQKLVGCGIG-----GDFYMAETPDVEVWGD 294  
 DB 43 MSFGIHRWRYTTEVSGARPMKVLGLAGTGLTAKFSLVGD-----KGEVWLD 95  
 OY 295 LSVNMISFALB-RSIGLKCAVEPADCTKINYPDNPSPDVIYSDRTILHIQDKALPRS 352  
 DB 96 INDSMLKGRTKLRDGIIVSNVSVQANAEALFPDNDHFDIITIAFGLNVTDKDALRS 155  
 OY 353 FYKMLKRGKVLISDYCKAGPSPEFAAVIKORGVLHDVK---EYGMQL-KDAGFVDV 408  
 DB 156 MVRVLPKPGKGLVLFESK-----PQHE-----VMKVKYDLVSFVLPVPGQLITKQADSYEY 207  
 OY 409 LAEDTEGFIVLRKLETVKEKXDFI-SDFSEEDYNDIVGG 450  
 DB 208 LAES-----IR-MHPDQDTL---KQMVVDAGFEQVDYTNNTDG 241

RESULT 6  
 GTOM ARATH STANDARD; PRT; 348 AA.  
 AC Q2SK1; Q9XIP9;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Tocopherol O-methyltransferase, chloroplast precursor (EC 2.1.1.95)  
 DE [Gmmt-tocopherol methyltransferase].  
 GN G-TMT OR AT1G64970 OR F13011.27.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosid II; Brassicales; Brassicaceae; Arabidopses.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX Shintani D.; DellaPenna D.;  
 RX MEDLINE=99069631; PubMed=9851934;  
 RT "Elevating the vitamin E content of plants through metabolic  
 engineering";  
 RL Science 283:2098-2100(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=21016719; PubMed=11130712;  
 RA White O.; Alonso J.; Altieri H.; Araujo R.; Bowman C.L.; Brooks S.Y.;  
 RA Buehler E.; Chan A.; Chao O.; Chen H.; Cheuk R.F.; Chin C.W.;  
 RA Chung M.K.; Conn L.; Conway A.B.; Conway T.H.; Dewar K.;  
 RA Dunn P.; Egu P.; Feldblum T.V.; Feng J.-D.; Pong B.; Fujii C.Y.;

RA Gill J.E.; Goldsmith A.D.; Haas B.; Hansen N.F.; Hughes B.; Huizar L.;  
 RA Hunter J.L.; Jenkins J.; Johnson-Hopson C.; Khan S.; Khaykin E.;  
 RA Kim C.J.; Koo H.L.; Kremenetskaia I.; Kurtz D.B.; Kvan A.; Lam B.;  
 RA Langin-Hooper S.; Lee A.; Lee J.M.; Lenz C.A.; Li J.H.; Li Y.-P.;  
 RA Lin X.; Liu S.X.; Liu Z.A.; Lueros J.S.; Maiti R.; Marziani A.;  
 RA Militscher J.; Miranda M.; Nguyen M.; Nierman W.C.; Osborne B.I.;  
 RA Pai G.; Peterson J.; Pham P.K.; Rizzo M.; Rooney T.; Rowley D.;  
 RA Sakano H.; Salzberg S.L.; Schwartz J.R.; Shinn P.; Southwick A.M.;  
 RA Sun H.; Tallon L.J.; Tambunga G.; Toriumi M.J.; Town C.D.;  
 RA Utterback T.; Van Aken S.; Vayberg M.; Vysotskaia V.S.; Walker M.;  
 RA Wu D.; Yu G.; Fraser C.M.; Venter J.C.; Davis R.W.;  
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
 RT thaliana";  
 RL Nature 408:816-820(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=22954850; PubMed=14593172;  
 RA Yamada K.; Lim J.; Dale J.M.; Chen H.; Shinn P.; Palm C.J.;  
 RA Southwick A.M.; Wu H.C.; Kim C.J.; Nguyen M.; Pham P.K.; Cheuk R.F.;  
 RA Karlin-Newmann G.; Liu S.X.; Lam B.; Sakano H.; Wu T.; Yu G.;  
 RA Miranda M.; Quach H.L.; Tripp M.; Chang C.H.; Lee J.M.; Toriumi M.J.;  
 RA Chan M.; Tang C.C.; Onodera C.S.; Deng J.M.; Akiyama K.; Carninci P.;  
 RA Arakawa T.; Banh J.; Banno F.; Bowser L.; Brooks S.Y.; Carninci P.;  
 RA Chao Q.; Choy N.; Enju A.; Goldsmith A.D.; Gurjal M.; Hansen N.F.;  
 RA Hayashizaki Y.; Johnson-Hopson C.; Hsuan V.W.; Iida K.; Karnes M.;  
 RA Khan S.; Koesema B.; Ishida J.; Jiang P.X.; Jones T.; Kawai J.;  
 RA Kamiya A.; Meyers C.; Nakajima M.; Narusaka M.; Seki M.; Sakurai T.;  
 RA Satou M.; Tanase R.; Vayberg M.; Wallender E.K.; Wong C.; Yamamura Y.;  
 RA Yuan S.; Shinozaki K.; Davis R.W.; Theologis A.; Ecker J.R.;  
 RT "Empirical analysis of transcriptional activity in the Arabidopsis  
 RT genome";  
 RL Science 302:842-846(2003).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC Brover V.; Troukhan M.; Alexandrov N.; Lu Y.-P.; Flavell R.;  
 RA Feldmann K.A.;  
 RT "Full-length cDNA from Arabidopsis thaliana";  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Methylates gamma- and delta-tocopherol to yield beta-  
 CC and alpha-tocopherol respectively.  
 CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + gamma-tocopherol =  
 CC S-adenosyl-L-homocysteine + alpha-tocopherol.  
 CC -!- PATHWAY: Vitamin E biosynthesis.  
 CC -!- SUBCELLULAR LOCATION: Chloroplast (By similarity).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AF104220; AAD02882.1; -  
 DR EMBL; AC006193; AAD38271.2; -  
 DR EMBL; AY049258; AAK83600.1; -  
 DR EMBL; AY090280; AAL90941.1; -  
 DR EMBL; AY087138; AAM64696.1; -  
 DR PIR; C96673; C96673.  
 DR InterPro; IPR001601; Methyltransf.  
 DR InterPro; IPR000051; SAM bind.  
 KW Methyltransferase; Transferase; Chloroplast; Transit peptide.  
 FT TRANSIT 1 50 CHLOROPLAST (POTENTIAL).  
 FT CHAIN 51 348 TOCOPHEROL O-METHYLTRANSFERASE.  
 FT DOMAIN 9 40 SER-RICH.  
 FT DOMAIN 123 232 SAM-BINDING.  
 FT CONFLICT 175 175 S -> A (IN REF. 2 AND 3).  
 FT CONFLICT 188 188 Q -> K (IN REF. 4).  
 SQ SEQUENCE 348 AA; 38091 MW; 47DB18A72E075F CRC64;

Query Match 6.9%; Score 173.5; DB 1; Length 348;  
 Best Local Similarity 27.2%; Pred. No. 9.1e-06;



EMBL; AP003134; BAB42564.1; --	Query Match	6.7%;	Score 167.5; DB 1; Length 241;
EMBL; AP004827; BAB95225.1; --	Best Local Similarity	23.0%;	Pred. No. 1.6e-05;
PIR; G89925; G89925.	Matches	56; Conservative	46; Mismatches 79; Indels 63; Gaps 8;
HAVAP; MF_01813; -- 1.			
InterPro; IPR001601; Methyltransf.			
InterPro; IPR000051; SAM_bind			
InterPro; IPR004034; Ubi7men_Metransf.			
InterPro; IPR004033; UbiE/COG5_Metrf.			
Pfam; PF01209; UbiE_methyltran; 1.			
PROSITE; PS01183; UBI1 1; FALSE NEG.			
PROSITE; PS01184; UBI2 2; 1.			
Menaquinone biosynthesis; Transferase; Methyltransferase;			
Complete proteome.			
SEQUENCE 241 AA; 27423 MW; 90BF4020ABCS54 CRC64;			
QY	197	NKGNQNIQSLWQVNSKDKGQRFELDTSQYKCHSILAYE--RVPGPGYVSTGGYETTK	255
DB	4	NKANKEGVHVRVFNISIKYD-----ALNNIISFEQHKVW-----RK	39
QY	255	EFVSKMLDKPQKQVLDVCGGIGGGDYMAETFDV----EVVGFDSVMNISPALERSIGL	310
DB	40	RVKMDGVKRGTKALDVCCGT--GDWTIISKAVGPTGEVTGIDFSENLEVGKKTASM	97
QY	311	KCAVEFEVADCTKINYPDNSFDVIYSKRDITLHIQKPALFRSYKWLKPGGKVLISDYCK	370
DB	98	E-NVRLVHGDMELPFEDNSFDVYTGFLGVNVPDVLVALKEMNRVLKPGGVV---CL	152
QY	371	KAGPP-----SPEPANYIKQGYDLHDVKEYGQMLKDG	404
DB	153	ETSOFTLVFVKQMYALYKFMPIFKGLFAKSKBEYEMLQQSTFTNPPGKBEIKRPFENG	212
QY	405	FVDV 408	
DB	213	FINV 216	

RESULT 8

UBIE\_VIBCH

ID	UBIE_VIBCH	STANDARD;	PRT;	260 AA.
AC	Q9KVQ6;			
DT	10-OCT-2003 (Rel. 42, Created)			
DT	10-OCT-2003 (Rel. 42, Last sequence update)			
DE	Ubiquinone/menaquinone biosynthesis (update)			
DE	Ubiquinone/menaquinone biosynthesis methyltransferase ubiE			
GN	[EC 2.1.1.-].			
GN	UBIE OR VC0083.			
OS	Vibrio cholerae.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;			
OC	Vibrionaceae; Vibrio.			
ON	NCBI_TaxID=666;			
OX	[1]			
RC	SEQUENCE FROM N.A.			
RC	STRAIN=El Tor N16961 / Serotype O1;			
RX	MEDLINE=J046833; PubMed=10952301;			
RA	Heidelberg J.C.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,			
RA	Dadson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,			
RA	Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson L.A.,			
RA	Smolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,			
RA	McDonald L., Utterback T., Fleischmann R.D., Nieman W.C., White O.,			
RA	Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,			
RA	Fraser C.M.;			
RT	"DNA sequence of both chromosomes of the cholera pathogen Vibrio			
RT	cholerae."			
RL	Nature 406:477-483(2000).			
CC	-1- FUNCTION: Methyltransferase required for the conversion of			
CC	dimethylmenaquinone (DMKH2) to menaquinone (MKH2) and the			
CC	conversion of 2-polyprenyl-6-methoxy-1,4-benzoquinol (ODMQH2) to			
CC	2-polyprenyl-3-methyl-6-methoxy-1,4-benzoquinol (DMQH2) (By			
CC	similarity).			



```

PT INIT MET 0 0 6.4% Score 161.5; DB 1; Length 381;
PT CONFLICT 1 1 S -> R (IN REP. 2);
PT CONFLICT 7 7 R -> G (IN REP. 2);
PT CONFLICT 24 24 S -> N (IN REP. 2);
PT CONFLICT 38 38 I -> T (IN REP. 2);
SQ SEQUENCE 381 AA; 43777 MW; 1F07B2240C7E08ADF CRC64;

Query Match 6.4%; Score 161.5; DB 1; Length 381;
Best Local Similarity 27.6%; Pred. No. 8.2e-05;
Matches 68; Conservative 26; Mismatches 91; Indels 61; Gaps 12;

QY 159 REPRTYKAFKCHLQDCGNSY-----ELSLSCKICAGYVRNK-----KNQOI- 204
DB 41 KNPDPFKRVLEGSU--GLGESYMGWECORLDMFPKSLVLAQENQLPHHFKOTLRIA 98
QY 205 -----SWLQK--VDSKDKGQRFIDT--SOYKNSILRYERYVCPG----- 244
DB 99 GARLNLQSKRQAVIGKEHYDLGNL--FSRMLDPPFQYSC-----ATWKDAD 145
QY 245 -VSTGGYETTKFVSMLDLKPQKQVLDVGGIGGDFYMAFTFVGVGFDLSVNMISFA 303
DB 146 NLESAQAQKLMKCEKQLKPKRMVLDIGCGWGLAHTMASNYDSVVGWTTISAPQQKMA 205
QY 304 LERSGLKCAVEFEVADCTKINYDP--NSPDVYSRDTLHIOCK--PALFRSFYKWLKP 359
DB 206 QRCCEGLDITLQ-----DYRLNDQFDRIVSGMPEHVGPKNYDTYFAVVDNLPK 258
QY 360 GQKVL 365
DB 259 EGIFLL 264

RESULT 10
UBIE_ECOL6 STANDARD; PRT; 251 AA.
AC Q8RFQ;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ubiquinone/menaquinone biosynthesis methyltransferase ubiE
(DB 2.1.1.-)
GN UBIE OR C4782.
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
[1]
SEQUENCE FROM N.A.
STRAIN=O6:HI / CFT073 / ATCC 700928;
MEDLINE=22388234; PubMed=12471157;
Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
"Extensive mosaic structure revealed by the complete genome sequence
of uropathogenic Escherichia coli";
Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
CC -!- FUNCTION: Methyltransferase required for the conversion of
dimethylmenaquinone (DMKH2) to menaquinone (MKH2) and the
conversion of 2-polyprenyl-6-methoxy-1,4-benzoquinol (DDMQH2) to
2-polyprenyl-3-methyl-6-methoxy-1,4-benzoquinol (DMQH2) (By
similarity).
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 2-polyprenyl-6-
methoxy-1,4-benzoquinol = S-adenosyl-L-homocysteine + 2-
polyprenyl-3-methyl-6-methoxy-1,4-benzoquinol.
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + demethylmenaquinol =
S-adenosyl-L-homocysteine + menaquinol.
CC -!- PATHWAY: Menaquinone biosynthesis; last step.
CC -!- PATHWAY: Ubiquinone biosynthesis.
CC -!- SIMILARITY: Belongs to the ubiE family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
or send an email to license@ebi.ac.uk).
EMBL; AE016770; AAN83215.1;
HAMAP; MF_01813; 1;
InterPro; IPR001601; Methyltransf.
InterPro; IPR000051; SAM_Dind.
InterPro; IPR004034; Ubi/men Methyltransf.
InterPro; IPR004033; Ubi/men Methyltransf.
Pfam; PF01209; Ubi methyltransf. 1.
PROSITE; PS01183; UBIIE 1; 1.
PROSITE; PS01184; UBIIE 2; 1.
Menaquinone biosynthesis; Ubiquinone biosynthesis; Transferase;
Methyltransferase; Complete proteome.
SQ SEQUENCE 251 AA; 28103 MW; B1EE209A997CA5E CRC64;

Query Match 6.4%; Score 160.5; DB 1; Length 251;
Best Local Similarity 28.0%; Pred. No. 5.7e-05;
Matches 60; Conservative 40; Mismatches 91; Indels 23; Gaps 9;

QY 245 VSTGGYETTKFVSMLDLKPQKQVLDVGGIGG--GDFYMAFTFVGVGFDLSVNMIS 301
DB 43 MSPGHRLWKKEFTIDCSGVRRGQTVLDLAGTGDLTAKEFSLVGETGKVLADINESMLK 102
QY 302 FALF--SSIGLKCAVEFEVADCTKINYDPNSPDVYSRDTLHIOCKPALFRSFYKWLKP 359
DB 103 MREKRLNIGVIGNVQNAALPPDNTFCITISFGLNVTDKALRSYRVLPK 162
QY 360 GQKVLSDYCKKAGPPSPFAVYKQKQVLDVGGIGG--GDFYMAFTFVGVGFDLSVNMIS 418
DB 163 GGRLLVLEFSPKIEPLSK--AY---DAYSFHVLPRIGSLVANDADSYRLAES-----I 212
QY 419 RVL--RKELETVEKEDVFISEEDYNDIVGG 450
DB 213 RMHPDQDTLTKMQD-----AGFESVDYNYLTAG 241

RESULT 11
UBIE_VIBPA STANDARD; PRT; 259 AA.
AC Q87TH4;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ubiquinone/menaquinone biosynthesis methyltransferase ubiE
(DB 2.1.1.-)
GN UBIE OR VP0095.
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
[1]
SEQUENCE FROM N.A.
STRAIN=RIMD 2210633 / Serotype O3:K6;
MEDLINE=22508454; PubMed=12620739;
Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
Iijima Y., Najima M., Nakano M., Yanashita A., Kubota Y., Kimura S.,
Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
"Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
distinct from that of V. cholerae";
Lancet 361:743-749 (2003).
CC -!- FUNCTION: Methyltransferase required for the conversion of
dimethylmenaquinone (DMKH2) to menaquinone (MKH2) and the
conversion of 2-polyprenyl-6-methoxy-1,4-benzoquinol (DDMQH2) to
2-polyprenyl-3-methyl-6-methoxy-1,4-benzoquinol (DMQH2) (By
similarity).
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 2-polyprenyl-6-
methoxy-1,4-benzoquinol = S-adenosyl-L-homocysteine + 2-
polyprenyl-3-methyl-6-methoxy-1,4-benzoquinol.
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + demethylmenaquinol =

```

[illegible]

```

DB 311 TVFRTSLRGKLVYRYSVQFLEK 332
|| : : | : : ||
RESULT 13
ID UBIE ECOLI STANDARD; PRT; 251 AA.
AC P27851;
DT 01-AUG-1992 (Rel. 23, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ubiquinone/menaquinone biosynthesis methyltransferase ubiE
(EC 2.1.1.-)
GN UBIE OR B3833 OR Z5355 OR EC54763 OR SF3911 OR S3843.
OS Escherichia coli.
OS Escherichia coli O157:H7, and
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 83334, 623;
[1]
RN SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12 / MG1655;
RX MEDLINE=92358234; PubMed=1379743;
RA Danielis D.L., Plunkett G. III, Burland V.D., Blattner F.R.;
RT "Analysis of the Escherichia coli genome: DNA sequence of the region
RT from 84.5 to 86.5 minutes.";
RL Science 257:771-778(1992).
[2]
RN SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-VIDES J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
[3]
RN SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Pofai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimallanta E.T., Potamouisis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
[4]
RN SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156211; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
[5]
RN SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Xian B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang S., Wen Y., Hou Y.,
RA Xu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).
[6]
RN SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.P., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786(2003).
[7]
RN CHARACTERIZATION, AND MUTANT UBIE401.
RC SPECIES=E.coli;
RX MEDLINE=97197541; PubMed=9045837;
RA Lee P.T., Hau A.Y., Ha H.T., Clarke C.F.;
RT "A C-methyltransferase involved in both ubiquinone and menaquinone
RT biosynthesis: isolation and identification of the Escherichia coli
RT ubiE gene.";
RL J. Bacteriol. 179:1748-1754(1997).
- i- FUNCTION: Methyltransferase required for the conversion of
- dimethylmenaquinone (DMKH2) to menaquinone (MKH2) and the
- conversion of 2-polyphenyl-6-methoxy-1,4-benzoquinol (DMQH2) to
- 2-polyphenyl-3-methyl-6-methoxy-1,4-benzoquinol (DMQH2).
- i- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 2-polyphenyl-6-
- methoxy-1,4-benzoquinol = S-adenosyl-L-homocysteine + 2-
- polyphenyl-3-methyl-6-methoxy-1,4-benzoquinol.
- i- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + demethylmenaquinol =
- S-adenosyl-L-homocysteine + menaquinol.
- i- PATHWAY: Menaquinone biosynthesis; last step.
- i- PATHWAY: Ubiquinone biosynthesis.
- i- SIMILARITY: Belongs to the ubiE family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
meddled and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
-----
EMBL; M87049; AAC7628.1; -
EMBL; AS000459; AAC76836.1; -
EMBL; AB005614; AAG59029.1; -
EMBL; AF002567; BAB38186.1; -
EMBL; AB015398; AAN45346.1; -
EMBL; AB016990; AAP18852.1; -
PIR; A86071; A86071.
PIR; B65188; B65188.
PIR; C98224; C98224.
ECGene; E01473; ubiE.
HAMAP; MF_01813; -; 1.
InterPro; IPR001601; Methyltransf.
InterPro; IPR000051; SAM bind.
InterPro; IPR004034; Ubi/men Methyltransf.
InterPro; IPR004033; UbiE/COQ5 Methyltransf.
Pfam; PF01209; UbiE methyltransf. 1.
PROSITE; PS01183; UBIE_1; 1.
PROSITE; PS01184; UBIE_2; 1.
Menaquinone biosynthesis; Ubiquinone biosynthesis; Transferase;
Menaquinone biosynthesis; Complete proteome.
Methyltransferase; Complete proteome.
MUTAGEN 142 142 L -> D: IN UBIE401, DEFECTIVE.
FT CONFLICT 101 101 L -> P (IN REF. 2).
SQ SEQUENCE 251 AA; 28073 MW; B1E9E209BDC68F5E CRC64;
Query Match 6.3%; Score 158.5; DB 1; Length 251;
Best Local Similarity 28.0%; Pred. NO. 8e-05;
Matches 60; Conservative 39; Mismatches 92; Indels 23; Gaps 10;
OY.. 245. VSTGGYETTKER VSMLLDKPQKVLGVGCGIGG--GDFYMAETFDVYVGFDSVNMIS 301
DB 43 MSFGIHLRWKRFIDCSGVRGQTVLDLGGTGTAKFSRLVGTGKVLADINESMLX 102

```

QY 302 FALF--RSLGKCAVEFEVADCTKINYPDNSFDVYSDTILHIQKPALEFSPYKWLKP 359  
 DB 103 MREKRLNVLGVNVEYVQANAEALPFPDFTFCITISFGLRNVTGDKALSKMTRVLKP 162  
 QY 360 GSKVLISDYCKAGKPPSPFAAYIKQRYDLHVDKVGQML-KDAGFVDVLAEDTEQFI 418  
 DB 163 GGRLLVLEFSPKPIEPLSK--AY---DAYSFHYLVPRIGSLVANDADSYRYLAES-----I 212  
 QY 419 RVLRKELETVRKEKDVFISSD--FSEEDYNDIVGG 450  
 DB 213 R-MHPDQDTLK-----ANMQDAGFSDVYNYLTAG 241

RESULT 14  
 UBIE\_VIBVU STANDARD; PRT; 260 AA.  
 AC Q8DDP9;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Ubiquinone/menaquinone biosynthesis methyltransferase ubiE  
 (EC 2.1.1.-).  
 GN UBIE OR V10909.  
 OS Vibrio vulnificus.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
 OC Vibrionaceae; Vibrrio.  
 OX NCBI\_TaxID=672;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CMCP6;  
 RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,  
 RA Choy H.B.;  
 RT "Complete genome sequence of *Vibrio vulnificus* CMCP6.";  
 RT Submitted (DEC 2002) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Methyltransferase required for the conversion of  
 CC dimethylmenaquinone (DMQH2) to menaquinone (MKH2) and the  
 CC conversion of 2-polyprenyl-6-methoxy-1,4-benzoquinol (DDMQH2) to  
 CC 2-polyprenyl-3-methyl-6-methoxy-1,4-benzoquinol (DMQH2) (By  
 CC similarity).  
 CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 2-polyprenyl-6-  
 CC methoxy-1,4-benzoquinol = S-adenosyl-L-homocysteine + 2-  
 CC polyprenyl-3-methyl-6-methoxy-1,4-benzoquinol.  
 CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + demethylmenaquinol =  
 CC S-adenosyl-L-homocysteine + menaquinol.  
 CC -!- PATHWAY: Menaquinone biosynthesis; last step.  
 CC -!- SIMILARITY: Belongs to the ubiE family.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; AS016800; AAC09411.1;  
 CC HAMAP; MF\_01813; -; 1.  
 CC InterPro; IPR001601; Methyltransf.  
 CC InterPro; IPR000051; SAM bind.  
 CC InterPro; IPR004034; Ubi/men.Metransf.  
 CC Pfam; PF01209; UbiE/COQ5\_Metrif.  
 CC PROSITE; PS01183; UBIE\_1; 1.  
 CC PROSITE; PS01184; UBIE\_2; 1.  
 CC Menaquinone biosynthesis; Ubiquinone biosynthesis; Transferase;  
 CC Methyltransferase; Complete proteome.  
 CC SEQUENCE 260 AA; 28971 MW; 81362EBDF89BE802 CRC64;

Query Match ..... 63% (score 158.5); DB 1; Length 260;  
 Best Local Similarity 26.1%; Pred. No. 8.4e-05;  
 Matches 57; Conservative 45; Mismatches 85; Indels 31; Gaps 11;

QY 245 VSTGYETTKEF--VSMLDLKPQKQVLDVCGGICGGGDFYMAETFDVEVVG-----FDLS 296  
 DB 52 MSGGTHRLWKKEFTDCSGARPGQVLDLGGGTG-----LTAKFS-RIVGKGHVILADIN 106  
 QY 297 VNMISFALE--RSLGKCAVEFEVADCTKINYPDNSFDVYSDTILHIQKPALEFSPY 354  
 DB 107 NSMLNVRGDKLRDSGIVGNVHYVQANAEELPPDDYDFDIITISFGLRNVTGDKALSKM 166  
 QY 355 KWLKGGKVLISDYCKKA-GPPSPFAAYIKQRYDLHVDKVGQML-KDAGFVDVLAED 412  
 DB 167 RVLRKELETVRKEKDVFISSD-----FSEEDYNDIVGG 450  
 QY 413 RTEQFIRVLRKELETVRKEKDVFISSDSEEDYNDIVGG 450  
 DB 221 -----IR-MHPDQDTLKGMQE--AGFENTSYNYLTGG 250

RESULT 15  
 UBIE\_NEIMA STANDARD; PRT; 245 AA.  
 AC Q8JVE3;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Ubiquinone/menaquinone biosynthesis methyltransferase ubiE  
 (EC 2.1.1.-).  
 GN UBIE OR NMA0956.  
 OS Neisseria meningitidis (serogroup A).  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;  
 OC Neisseriaceae; Neisseria.  
 OX NCBI\_TaxID=65699;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=22491 / Serogroup A / Serotype 4A;  
 RX MEDLINE=2022556; PubMed=10761919;  
 RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,  
 RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,  
 RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,  
 RA Jagals K., Leather S., Moule S., Mungall K., Quail M.A.,  
 RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,  
 RA Whitehead S., Spratt B.G., Barrall B.G.;  
 RT "Complete DNA sequence of a serogroup A strain of *Neisseria*  
 RT *meningitidis* 22491.";  
 RL Nature 404:502-506(2000).  
 CC -!- FUNCTION: Methyltransferase required for the conversion of  
 CC dimethylmenaquinone (DMQH2) to menaquinone (MKH2) and the  
 CC conversion of 2-polyprenyl-6-methoxy-1,4-benzoquinol (DDMQH2) to  
 CC 2-polyprenyl-3-methyl-6-methoxy-1,4-benzoquinol (DMQH2) (By  
 CC similarity).  
 CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 2-polyprenyl-6-  
 CC methoxy-1,4-benzoquinol = S-adenosyl-L-homocysteine + 2-  
 CC polyprenyl-3-methyl-6-methoxy-1,4-benzoquinol.  
 CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + demethylmenaquinol =  
 CC S-adenosyl-L-homocysteine + menaquinol.  
 CC -!- PATHWAY: Menaquinone biosynthesis; last step.  
 CC -!- SIMILARITY: Belongs to the ubiE family.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; AL162754; CAB84226.1; -;  
 CC FIR; C81942; C81942.  
 CC HAMAP; MF\_01813; -; 1.  
 CC InterPro; IPR001601; Methyltransf.  
 CC InterPro; IPR000051; SAM bind.  
 CC InterPro; IPR004034; Ubi/men.Metransf.  
 CC InterPro; IPR004033; UbiE/COQ5\_Metrif.

DR Pfam: PF01209; Ubic methyltran; 1.  
DR PROSITE: PS01183; UBI1; 1.  
DR PROSITE: PS01184; UBI2; 1.  
KW Ubiquinone biosynthesis; Ubiquinone biosynthesis; Transferase;  
KW Methyltransferase; Complete proteome.  
SQ SEQUENCE 245 AA; 27337 MW; 256E95211E7551C7 CRC64;

Query Match 6.2%; Score 157; DB 1; Length 245;  
Best Local Similarity 25.8%; Pred. No. 0.0001;  
Matches 55; Conservative 45; Mismatches 89; Indels 24; Gaps 8;

QY	245	VSTGGYETTKBP-VSMLELKEGQKVLGVGGIGGGDFYMAETP--DVEYVGPDLSEVNMIS	301
DB	40	MSAGLHRVWKHPTINTAHLKKGDKVLDIAGTGDLSKGWAKRVGKEGVMLTDINSSMLT	99
QY	302	PALEISIGLKCAVEPEVADCTKINYPNPSFDVIYSRDTILHIQDKPALFRSFKYKWKPGG	361
DB	100	VGRDRLNEGMLFVSLDAEKLFPDNYFNLVSVAFLNTHKDAALKENYRVLKPGG	159
QY	362	KVLISDYCKGAGPPSPFAAYIKORYDLHDVK---EYGM-L-KDAGFVDVLAEDRTQF	417
DB	160	TLVLVLEFSKIYKP-----LEGAYDFTSFKLLPANGRLIAKDAESYQYLAES-----	205
QY	418	IRVLKRLKLTVEKKEKDVFIISDFSEEDYNDIVGG	450
DB	206	IR-MHPDQETLAKM--MLDAGFDSVDYHNKSG	235

Search completed: July 26, 2004, 13:14:33  
Job time : 14 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2004, 13:11:41 ; Search time 41 Seconds  
(without alignments)  
3640.003 Million cell updates/sec

Title: US-10-031-331B-40

Perfect score: 2515

Sequence: 1 HTVDLTIEMMLDQASDL.....KLRRATGQRWGLPVAKK 473

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:\*

- 1: sp arches:\*
- 2: sp bacteria:\*
- 3: sp fungi:\*
- 4: sp human:\*
- 5: sp invertebrate:\*
- 6: sp mammal:\*
- 7: sp mhc:\*
- 8: sp organelle:\*
- 9: sp phage:\*
- 10: sp plant:\*
- 11: sp rodent:\*
- 12: sp virus:\*
- 13: sp vertebrate:\*
- 14: sp unclassified:\*
- 15: sp virus:\*
- 16: sp bacteriap:\*
- 17: sp archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2515	100.0	494	10 Q85287	Q85287 suada Japo
2	2063	82.0	491	10 Q9AXH3	Q9AXH3 lycopersico
3	2044.5	81.3	493	10 Q84SA4	Q84SA4 aster tripo
4	2009	79.9	491	10 Q7XJZ2	Q7XJZ2 brassica na
5	1968	78.3	499	10 Q8LJ10	Q8LJ10 oryza sativ
6	1918.5	76.3	498	10 Q8VXK1	Q8VXK1 triticum ae
7	1515	60.2	376	10 Q8L7A8	Q8L7A8 arabidopsis
8	516.5	20.5	437	5 Q22933	Q22933 caenorhabdi
9	476	18.9	266	5 Q81DQ9	Q81DQ9 plasmodium
10	413	16.4	460	5 Q86NE3	Q86NE3 caenorhabdi
11	413	16.4	475	5 Q23552	Q23552 caenorhabdi
12	413	16.4	484	5 Q95PW7	Q95PW7 caenorhabdi
13	288.5	11.5	264	16 Q98VFB	Q98VFB rhizobium 1
14	238	9.5	351	5 Q8IFX3	Q8IFX3 caenorhabdi
15	229.5	9.1	270	2 Q83X76	Q83X76 streptomyce
16	225.5	9.0	276	2 Q846Y0	Q846Y0 streptomyce

17	222	8.8	565	2	Q9KJ20	Q9KJ20 actinopolya
18	218.5	8.7	283	2	Q9X5Q9	Q9X5Q9 streptomyce
19	217.5	8.6	363	10	Q41587	Q41587 triticum ae
20	217.5	8.6	363	10	Q41586	Q41586 triticum ae
21	214	8.5	280	16	Q81W13	Q81W13 anabaena sp
22	211.5	8.4	346	10	Q82720	Q82720 nicotiana t
23	210.5	8.4	317	16	Q55609	Q55609 synecocyst
24	209.5	8.3	272	2	Q8K152	Q8K152 nocardia ae
25	206.5	8.2	273	2	Q8K152	Q8K152 nocardia ae
26	206.5	8.2	283	2	Q8K294	Q8K294 nocardia ae
27	203	8.1	290	16	Q8KBN7	Q8KBN7 chlorobium
28	200.5	8.0	279	2	Q9KJ21	Q9KJ21 ectothiorho
29	200	8.0	278	2	Q9EY12	Q9EY12 streptomyce
30	199.5	7.9	349	10	Q82434	Q82434 nicotiana t
31	199	7.9	344	10	Q49215	Q49215 zea mays (m
32	198	7.9	344	10	Q93852	Q93852 zea mays (m
33	197	7.8	276	2	Q83WG2	Q83WG2 streptomyce
34	196	7.8	330	10	Q84M50	Q84M50 oryza sativ
35	194.5	7.7	283	2	Q9ALN3	Q9ALN3 saccharopol
36	193	7.7	367	10	Q43445	Q43445 glycine max
37	190	7.6	349	10	Q82426	Q82426 oryza sativ
38	189.5	7.5	280	2	Q83WF7	Q83WF7 streptomyce
39	187.5	7.5	330	16	Q8YV60	Q8YV60 anabaena sp
40	186.5	7.4	335	10	Q9LM02	Q9LM02 arabidopsis
41	185.5	7.4	227	17	Q57965	Q57965 pyrococcus
42	184.5	7.3	318	16	P74388	P74388 synecocyst
43	184.5	7.3	336	10	Q8LKW1	Q8LKW1 arabidopsis
44	184	7.3	277	2	Q83WC3	Q83WC3 aphanothese
45	183.5	7.3	387	16	Q9PNB2	Q9PNB2 campylobact

#### ALIGNMENTS

#### RESULT 1

Q85287	PRELIMINARY; PRT; 494 AA.
ID	Q85287
AC	Q85287; TREMBLrel. 24, Created)
DT	01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT	01-JUN-2003 (TREMBLrel. 25, Last annotation update)
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE	Phosphoethanolamine N-methyltransferase.
GN	PEAMT
OS	Suadea japonica.
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC	Caryophyllales; Amaranthaceae; Suadea.
OX	NCBI_TaxID=90346;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Yamada A., Nozawa G.T., Tanimoto S., Ozeki Y.;
RT	"Glycinebetaine synthesis in Suadea japonica."
RL	Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AB080186; BAC57432.1;
DR	GO; GO:0008755; F:S-adenosylmethionine-dependent methyltransf. . . ; IEA.
DR	GO; GO:0016740; F:transferase activity; IEA.
DR	InterPro; IPR001601; Methyltransf.
DR	InterPro; IPR00051; SAM bind.
DR	InterPro; IPR004033; UbiE/COQ5 Metrf.
DR	Pfam; PF01209; Ubie methyltran; 1.
KW	Transferase; Methyltransferase.
SQ	SEQUENCE 494 AA; 56562 MW; BC613P9097BD3AE3 CRC64;

Query Match 100.0%; Score 2515; DB 10; Length 494;  
Best Local Similarity 100.0%; Pred. No. 1.9e-188;  
Matches 473; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HTVDLTIEMMLDQASDLQKERPELTLMLPPEKCLLELCAGIGRTGSLAKAQV 60

22 HTVDLTIEMMLDQASDLQKERPELTLMLPPEKCLLELCAGIGRTGSLAKAQV 81

Qy 61 IALDPIESAIKQNVINGKVKFKCVDVTSPTLSPPHSLDVTFSNMLWLSDEVE 120

Db 82 IALDPIESAIKKEVINGHYKVKFMCADVTSPILSPFHSLDVIIPSNMLLWYLSDEVE 141  
QY 121 NLVERMLKWKPGGYIPFSCFHQGDHRSNTHYRPRFTYKPKCEHLOQSGNS 180  
Db 142 NLVERMLKWKPGGYIPFSCFHQGDHRSNTHYRPRFTYKPKCEHLOQSGNS 201  
QY 181 YELSLSCCKICIGAYVRNKKQNIQISWLQKVDSDKDKGQFRLDTSQYKCNILRYRVP 240  
Db 202 YELSLSCCKICIGAYVRNKKQNIQISWLQKVDSDKDKGQFRLDTSQYKCNILRYRVP 261  
QY 241 GPGVSTGGYETTKFVSMGLDKPGQKVDGCGIGGGDFYMAETFDVEVVGFDLSVNM 300  
Db 262 GPGVSTGGYETTKFVSMGLDKPGQKVDGCGIGGGDFYMAETFDVEVVGFDLSVNM 321  
QY 301 SPALERSIGLKCAREFVADCTKINPDNSFDVIYSRDTILHIDKPALEFSPYKMLKPG 360  
Db 322 SPALERSIGLKCAREFVADCTKINPDNSFDVIYSRDTILHIDKPALEFSPYKMLKPG 381  
QY 361 KGLVLSIDYCKKAGPPSPFAAYIKQGYDLHDVKEYGQMLKADGFDVLAEDRTQFIRV 420  
Db 382 KGLVLSIDYCKKAGPPSPFAAYIKQGYDLHDVKEYGQMLKADGFDVLAEDRTQFIRV 441  
QY 421 LRKELETVEKQVPIISDFSEEDYNDIVGQNDKLRATKGEORWGLFVAKKK 473  
Db 442 LRKELETVEKQVPIISDFSEEDYNDIVGQNDKLRATKGEORWGLFVAKKK 494

RESULT 2  
Q9AMX3 PRELIMINARY; PRT; 491 AA.  
ID Q9AMX3  
AC Q9AMX3; 17, Created)  
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)  
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)  
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
DE Phosphoethanolamine N-methyltransferase.  
OS Lycopersicon esculentum (tomato).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC Lamiales; Solanales; Solanaceae; Solanum.  
CX NCBI\_TaxID=4081;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Parani M., Parida A.;  
RT "Characterization of a cDNA for phosphoethanolamine N-methyltransferase"; to the EMBL/GenBank/DBJ databases.  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
DR GO; GO:0005344; F:oxyanion transporter activity; IEA.  
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransferase.  
DR GO; GO:0016740; P:transferase activity; IEA.  
DR GO; GO:0006810; P:transferase activity; IEA.  
DR InterPro; IPR000896; Hemocyanin.  
DR InterPro; IPR001601; SAM bind.  
DR InterPro; IPR000051; SAM bind.  
DR PROSITE; PS00210; HEMOCYANIN 2; 1.  
KW Methyltransferase; Transferase.  
SQ SEQUENCE 491 AA; 55947 MW; D3GECDD8512733E7 CRC64;

Query Match 82.0%; Score 2063; DB 10; Length 491;  
Best Local Similarity 80.1%; Pred. No. 4.7e-153;  
Matches 379; Conservative 46; Mismatches 48; Indels 0; Gaps 0;

QY 1 HTVDTLTIAMMLDSQASDLDEKPEPILSMPLPGLGKLELGLAGIGRFTGELAEKAGQV 60  
Db 19 HTAELTVAMMLDSKAADLDEKPEPILSMPLPGLGKLELGLAGIGRFTGELAEKAGQV 78  
QY 61 IALDPIESAIKKEVINGHYKVKFMCADVTSPILSPFHSLDVIIPSNMLLWYLSDEVE 120  
Db 79 IALDPIESAIKKEVINGHYKVKFMCADVTSPILSPFHSLDVIIPSNMLLWYLSDEVE 138  
QY 121 NLVERMLKWKPGGYIPFSCFHQGDHRSNTHYRPRFTYKPKCEHLOQSGNS 180  
Db 139 ALVERMWLWVGCHIFFRSCFHQGDHRSNTHYRPRFTYKPKCEHLOQSGNS 198

QY 181 YELSLSCCKICIGAYVRNKKQNIQISWLQKVDSDKDKGQFRLDTSQYKCNILRYRVP 240  
Db 199 FELSLAGCKICIGAYVRNKKQNIQISWLQKVDSDKDKGQFRLDTSQYKCNILRYRVP 258  
QY 241 GPGVSTGGYETTKFVSMGLDKPGQKVDGCGIGGGDFYMAETFDVEVVGFDLSVNM 300  
Db 259 GPGVSTGGYETTKFVSMGLDKPGQKVDGCGIGGGDFYMAETFDVEVVGFDLSVNM 318  
QY 301 SPALERSIGLKCAREFVADCTKINPDNSFDVIYSRDTILHIDKPALEFSPYKMLKPG 360  
Db 319 SPALERSIGLKCAREFVADCTKINPDNSFDVIYSRDTILHIDKPALEFSPYKMLKPG 378  
QY 361 KGLVLSIDYCKKAGPPSPFAAYIKQGYDLHDVKEYGQMLKADGFDVLAEDRTQFIRV 420  
Db 379 KGLVLSIDYCKKAGPPSPFAAYIKQGYDLHDVKEYGQMLKADGFDVLAEDRTQFIRV 438  
QY 421 LRKELETVEKQVPIISDFSEEDYNDIVGQNDKLRATKGEORWGLFVAKKK 473  
Db 439 LRKELETVEKQVPIISDFSEEDYNDIVGQNDKLRATKGEORWGLFVAKKK 491

RESULT 3  
Q84SA4 PRELIMINARY; PRT; 493 AA.  
ID Q84SA4  
AC Q84SA4; 24, Created)  
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)  
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE Phosphoethanolamine N-methyltransferase.  
OS Aster tripolium (sea aster).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC Campanulids; Asterales; Asteraceae; Asteroideae; Astereae; Aster.  
CX NCBI\_TaxID=74787;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Takeda M., Uno Y., Kanechi M., Inagaki N.;  
RT "Analysis of nine cDNAs for salt-inducible genes in the halophyte sea aster."; to the EMBL/GenBank/DBJ databases.  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB090883; BAC57960.1; -;  
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransferase.  
DR GO; GO:0016740; P:transferase activity; IEA.  
DR InterPro; IPR001601; Methyltransf.  
DR InterPro; IPR000051; SAM bind.  
DR Methyltransferase; Transferase.  
KW Methyltransferase; Transferase.  
SQ SEQUENCE 493 AA; 56002 MW; 9F2C7369192B6DA5 CRC64;

Query Match 81.3%; Score 2044.5; DB 10; Length 493;  
Best Local Similarity 78.9%; Pred. No. 1.3e-151;  
Matches 374; Conservative 51; Mismatches 48; Indels 1; Gaps 1;

QY 1 HTVDTLTIAMMLDSQASDLDEKPEPILSMPLPGLGKLELGLAGIGRFTGELAEKAGQV 60  
Db 20 HSDVLTAVESMLDSMASDLDEKPEPILSMPLPGLGKLELGLAGIGRFTGELAEKAGQV 79  
QY 61 IALDPIESAIKKEVINGHYKVKFMCADVTSPILSPFHSLDVIIPSNMLLWYLSDEVE 120  
Db 80 IALDPIESAIKKEVINGHYKVKFMCADVTSPILSPFHSLDVIIPSNMLLWYLSDEVE 139  
QY 121 NLVERMLKWKPGGYIPFSCFHQGDHRSNTHYRPRFTYKPKCEHLOQSGNS 180  
Db 140 DIABRFLKWKVGGHIFRSCFHQGDHRSNTHYRPRFTYKPKCEHLOQSGNS 199  
QY 181 YELSLSCCKICIGAYVRNKKQNIQISWLQKVDSDKDKGQFRLDTSQYKCNILRYRVP 239  
Db 200 YELSLSCCKICIGAYVRNKKQNIQISWLQKVDSDKDKGQFRLDTSQYKCNILRYRVP 259  
QY 240 GPGVSTGGYETTKFVSMGLDKPGQKVDGCGIGGGDFYMAETFDVEVVGFDLSVNM 299  
Db 260 GPGVSTGGYETTKFVSMGLDKPGQKVDGCGIGGGDFYMAETFDVEVVGFDLSVNM 319

QY 300 ISFALERSIGLKCAVEFEVADCTKINYPDNSFDVIYSRDTILHIQDKPALFRSFKWLPK 359  
 Db 320 IAFALERSIGLQCSVEFEVADCTKSPDNSFDVIYSRDTILHIQDKPALFRFTWLPK 379  
 QY 360 GKVLISDYCKAGPPSPFAAYIKQGYDLHDVKEYGOMLKDAGFVDVLAEDRTQPIR 419  
 Db 380 GKVLISDYCKAGPPSPFAAYIKQGYDLHDVKEYGOMLKDAGFVDVLAEDRTQPIR 439  
 QY 420 VLKSELTVEKEKOVFISDFSEEDYNDIVGWNDRKLRRTAKGEQWGLFVAKK 473  
 Db 440 VLKSELTVEKEKOVFISDFSEEDYNDIVGWNDRKLRRTAKGEQWGLFVAKK 493

RESULT 4  
 Q7XJ22 PRELIMINARY; PRT; 491 AA.  
 ID Q7XJ22  
 AC Q7XJ22  
 DT 01-OCT-2003 (TREMBlrel. 25, Created)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)  
 DE Phosphothanolamine N-methyltransferase.  
 OS Brassica napus (Rape).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids1; Brassicales; Brassicaceae; Brassica.  
 OX NCBI\_TaxID=3708;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Ye C., Li J., Yang J., Wang B.;  
 RA "Plant gene cloning and expression analysis under different  
 RT stresses";  
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY319479; AAP3582.1; -;  
 KW Methyltransferase; Transferase.  
 SQ SEQUENCE 491 AA; 55398 MW; A08C3318737031EE CRC64;

Query Match 79.9%; Score 2009; DB 10; Length 491;  
 Best Local Similarity 78.0%; Pred. No. 7.9e-149;  
 Matches 368; Conservative 52; Mismatches 52; Indels 0; Gaps 0;

QY 1 HTVDITIAMLDLSDQASDLKDEPERPEILSMPLPEGLKLELGAGIGRTFGLAKAGOV 60  
 Db 19 HSADLTVEAMLDLSDQASDLKDEPERPEILSMPLPEGLKLELGAGIGRTFGLAKAGEL 78  
 QY 61 IALDFIESAIKKNVINGHYKVKFCADVTSPTLSPPPHSLDVIPIFNMLLVLSDEIVE 120  
 Db 79 IALDFISVINKNSVINGHYKVKFCADVTSPTLSPPPHSLDVIPIFNMLLVLSDEIVE 138  
 QY 121 NLVVERMLKWLKPGYIFPSCFHSQSDHKKSNPTHYREPRPYTKAPKECHLODGSNS 180  
 Db 139 LLVERMVGIKVGVIYFSCFHSQSDHKKSNPTHYREPRPYTKAPKECHLODGSNS 198  
 QY 181 YELSLSCCKICGAYVRNKKQNOISWLWQVDSKDKGFORPLDTQSYKNSILYERVP 240  
 Db 199 FELSMIGKCIKICGAYVRNKKQNOISWLWQVDSKDKGFORPLDTQSYKNSILYERVP 258  
 QY 241 GPGYVSTGGYETTKFVSMGLDKPGQKLVGCGIGGDFYMAETPDVSVVGLDLSVNI 300  
 Db 259 GGGFVSTGGIETTKFVSMGLDKPGQKLVGCGIGGDFYMAETPDVSVVGLDLSVNI 318  
 QY 301 SPALERSIGLKCAVEFEVADCTKINYPDNSFDVIYSRDTILHIQDKPALFRSFKWLPK 360  
 Db 319 SPALERAIGLNCVSEFEVADCTKHYKSPNSKDVYISRDTILHIQDKPALFRFTWLPK 378  
 QY 361 GKVLISDYCKAGPPSPFAAYIKQGYDLHDVKEYGOMLKDAGFVDVLAEDRTQPIR 420  
 Db 379 GKVLISDYCKAGPPSPFAAYIKQGYDLHDVKEYGOMLKDAGFVDVLAEDRTQPIR 438  
 QY 421 LRKELETVEKEKOVFISDFSEEDYNDIVGWNDRKLRRTAKGEQWGLFVAKK 472  
 Db 439 LRRELVEKEKKEEFISDFSEEDYNDIVGWNDRKLRRTAKGEQWGLFVAKK 490

RESULT 5  
 Q8LJ10 PRELIMINARY; PRT; 499 AA.  
 ID Q8LJ10  
 AC Q8LJ10  
 DT 01-OCT-2002 (TREMBlrel. 22, Created)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
 DE Putative phosphothanolamine methyltransferase.  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. Nipponbare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC  
 RL clone: P0431H09";  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP003248; BAC10708.1; -;  
 DR Gramene; Q8LJ10; -;  
 DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. .; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR InterPro; IPR001601; Methyltransf.  
 DR InterPro; IPR00051; SAM bind.  
 KW Transferase; Methyltransferase.  
 SQ SEQUENCE 499 AA; 56786 MW; ADDF3D04862D18BE CRC64;

Query Match 78.3%; Score 1968; DB 10; Length 499;  
 Best Local Similarity 76.1%; Pred. No. 1.3e-145;  
 Matches 359; Conservative 52; Mismatches 61; Indels 0; Gaps 0;

QY 1 HTVDITIAMLDLSDQASDLKDEPERPEILSMPLPEGLKLELGAGIGRTFGLAKAGOV 60  
 Db 28 HSKDLTVEAMLDLSDQASDLKDEPERPEILSMPLPEGLKLELGAGIGRTFGLAKAGV 87  
 QY 61 IALDFIESAIKKNVINGHYKVKFCADVTSPTLSPPPHSLDVIPIFNMLLVLSDEIVE 120  
 Db 88 LAMDFIESVINKNSVINGHYKVKFCADVTSPTLSPPPHSLDVIPIFNMLLVLSDEIVE 147  
 QY 121 NLVVERMLKWLKPGYIFPSCFHSQSDHKKSNPTHYREPRPYTKAPKECHLODGSNS 180  
 Db 148 KLVKRWMLKVGVIYFSCFHSQSDHKKSNPTHYREPRPYTKAPKECHLODGSNS 207  
 QY 181 YELSLSCCKICGAYVRNKKQNOISWLWQVDSKDKGFORPLDTQSYKNSILYERVP 240  
 Db 208 FELSLVLTCKVGVKVKSNQNOICMLWQVDSKDKGFORPLDTQSYKNSILYERVP 267  
 QY 241 GPGYVSTGGYETTKFVSMGLDKPGQKLVGCGIGGDFYMAETPDVSVVGLDLSVNI 300  
 Db 268 GGGFVSTGGIETTKFVSMGLDKPGQKLVGCGIGGDFYMAETPDVSVVGLDLSVNI 327  
 QY 301 SPALERSIGLKCAVEFEVADCTKINYPDNSFDVIYSRDTILHIQDKPALFRSFKWLPK 360  
 Db 328 SPALERAIGRCSVEFEVADCTKTKYKSPNSKDVYISRDTILHIQDKPALFRFTWLPK 387  
 QY 361 GKVLISDYCKAGPPSPFAAYIKQGYDLHDVKEYGOMLKDAGFVDVLAEDRTQPIR 420  
 Db 388 GKVLISDYCKAGPPSPFAAYIKQGYDLHDVKEYGOMLKDAGFVDVLAEDRTQPIR 447  
 QY 421 LRKELETVEKEKOVFISDFSEEDYNDIVGWNDRKLRRTAKGEQWGLFVAKK 472  
 Db 448 LRRELVEKEKKEEFISDFSEEDYNDIVGWNDRKLRRTAKGEQWGLFVAKK 499

RESULT 6  
 Q8VYX1 PRELIMINARY; PRT; 496 AA.  
 ID Q8VYX1  
 AC Q8VYX1  
 DT 01-MAR-2002 (TREMBlrel. 20, Created)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)

DN 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
DE Phosphoethanolamine methyltransferase.  
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae;  
OC Triticaceae; Triticum.  
OX NCBI\_TaxID=4565;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Prenette Charon J.-B., Breton G., Danyluk J., Muzac I., Ibrahim R.,  
SA Sarhan P.;  
RT "Molecular and biochemical characterization of a cold regulated  
phosphoethanolamine methyltransferase from wheat."  
RL Submitted (DSC-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY065971; AAL0895.1;  
DR GO; GO:0008757; P:S-adenosylmethionine-dependent methyltransf. . . IEA.  
DR GO; GO:0016740; P:transferase activity; IEA.  
DR InterPro; IPR001601; Methyltransf.  
DR InterPro; IPR000051; SAM bind.  
KW Transferase; Methyltransferase.  
SQ SEQUENCE 498 AA; 56858 MW; 37BB7134E2DEA148 CRC64;  
  
Query Match 76.34; Score 1918.5; DB 10; Length 498;  
Best Local Similarity 73.84; Pred. No. 9.9e-142;  
Matches 349; Conservative 59; Mismatches 64; Indels 1; Gaps 1;  
  
QY 1 HTVDLTTAMMLDSQADLDKEERPEILSMPLPESGKCLLELGAGIGRTGELAKAGOV 60  
DB 26 HSKDLTVSMMLDSRAKDLKBERPRVLAIPYAGKTVLELGAGIGRTGELAKAGV 85  
QY 61 IALDPFESAIKKNVEINGH-YKNVKNFCADVTSPTLSFPFPHSLDVIFSNWLLMYLSDSEV 119  
DB 86 IALDPFDSVIVKKNBEINGDIYKNITFCADVTSPELKIEDNSVDIVFSNWLMLYLNDEV 145  
QY 120 ENLVERMLKWLPGGYIFPESCFHQSDHKNKSNPTHYEPFYTAKFKECHLQDGSN 179  
DB 146 EKLIGRIVKWLPGGYIFPESCFHQSDHKNKSNPTHYEPFYTAKFKECHLQDGSN 205  
QY 180 SYELSLSCCKTGAYVRNKKQNIQISMLWQVDSKDKGQFRLDTSQYKNSILYERVF 239  
DB 206 SPELSLVTSKCIKAYVRNKKQNIQISMLWQVDSKDKGQFRLDTSQYKNSILYERVF 265  
QY 240 FGRGYVSTGYETTKFYSMLDLKPGQKVLVGGGIGGDFYMAETFDVVEVVGFDLSVM 299  
DB 266 FGRGYVSTGYETTKFYSMLDLKPGQKVLVGGGIGGDFYMAETFDVVEVVGFDLSVM 325  
QY 300 ISFALERSGLKCAVEFEVADCTKINYPDSNDFVYISRDTHLQDKPALFRSFKWLKP 359  
DB 326 VSFALERSGLKCAVEFEVADCTKINYPDSNDFVYISRDTHLQDKPALFRSFKWLKP 385  
QY 360 GGVKLVSDYCKKAGPPSPFAAYIKQGYDLHDVKEYGQMLKQAGFVDVLAEORTQFTR 419  
DB 386 GGVKLVSDYCKRSPGTPSPFAAYIKQGYDLHDVKEYGQMLKQAGFVDVLAEORTQFTR 445  
QY 420 VLKELTVEKEKQVPSPESSDYNDIVGWNCKLRTAKGKGLFVAKK 472  
DB 446 VLKELTVEKEKQVPSPESSDYNDIVGWNCKLRTAKGKGLFVAKK 498  
  
RESULT 7  
Q8L7A8 PRELIMINARY; PRT; 376 AA.  
ID Q8L7A8  
AC Q8L7A8  
DT 01-OCT-2002 (TRENBLrel. 22, Created)  
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)  
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
DE Phosphoethanolamine N-methyltransferase, putative.  
GN ATIG73600.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsie.  
OX NCBI\_TaxID=3702;

RN SEQUENCE FROM N.A.  
RA Southwick A., Nguyen M., Tripp M., Palm C.J., Jones T., Wu T.,  
RA Carninci P., Chen H., Cheuk R., Chan M.M., Chang C.H., Dale J.M.,  
RA Deng J.M., Hayashizaki Y., Hsuan V.W., Lee J.M., Ishida J., Kamiya A.,  
RA Kawai J., Kim C.J., Narusaka M., Quach H.L., Sakurai T., Satou M.,  
RA Seki M., Shimizu P., Tang C.C., Toroumi M., Wallander E.K., Wong C.,  
RA Wu H.C., Yamada K., Yu G., Yuan S., Shinozaki K., Ecker J.,  
RA Theologis A., Davis R.W.;  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY136372; AA97038.1;  
DR GO; GO:0008757; P:S-adenosylmethionine-dependent methyltransf. . . IEA.  
DR GO; GO:0016740; P:transferase activity; IEA.  
DR InterPro; IPR001601; Methyltransf.  
DR InterPro; IPR000051; SAM bind.  
KW Transferase; Methyltransferase.  
SQ SEQUENCE 376 AA; 42811 MW; 2B4F6BE2B112B3ED CRC64;  
  
Query Match 60.24; Score 1515; DB 10; Length 376;  
Best Local Similarity 81.14; Pred. No. 2.7e-110;  
Matches 279; Conservative 31; Mismatches 34; Indels 0; Gaps 0;  
  
QY 1 HTVDLTTAMMLDSQADLDKEERPEILSMPLPESGKCLLELGAGIGRTGELAKAGOV 60  
DB 32 HSKDLTVSMMLDSRAKDLKBERPRVLAIPYAGKTVLELGAGIGRTGELAKAGV 91  
QY 61 IALDPFESAIKKNVEINGH-YKNVKNFCADVTSPTLSFPFPHSLDVIFSNWLLMYLSDSEV 120  
DB 92 IALDPFESVIVKKNBEINGDIYKNITFCADVTSPELKIEDNSVDIVFSNWLMLYLNDEV 151  
QY 121 NLVERMLKWLPGGYIFPESCFHQSDHKNKSNPTHYEPFYTAKFKECHLQDGSN 180  
DB 152 DLAKGGLQWTKVGGYIFPESCFHQSDHKNKSNPTHYEPFYTAKFKECHLQDGSN 211  
QY 181 YELSLSCCKTGAYVRNKKQNIQISMLWQVDSKDKGQFRLDTSQYKNSILYERVF 240  
DB 212 YELSLVSCCKTGAYVRNKKQNIQISMLWQVDSKDKGQFRLDTSQYKNSILYERVF 271  
QY 241 FGRGYVSTGYETTKFYSMLDLKPGQKVLVGGGIGGDFYMAETFDVVEVVGFDLSVM 300  
DB 272 GGVKLVSDYCKKAGPPSPFAAYIKQGYDLHDVKEYGQMLKQAGFVDVLAEORTQFTR 331  
QY 301 ISFALERSGLKCAVEFEVADCTKINYPDSNDFVYISRDTHLQDKPALFRSFKWLKP 344  
DB 332 ISFALERSGLKCAVEFEVADCTKINYPDSNDFVYISRDTHLQDKPALFRSFKWLKP 375  
  
RESULT 8  
Q22993 PRELIMINARY; PRT; 437 AA.  
ID Q22993  
AC Q22993  
DT 01-NOV-1996 (TRENBLrel. 01, Created)  
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
GN F54D11.1.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN=Bristol N2;  
RX MEDLINE=99059613; PubMed=9851916;  
RA None;  
RT "Genome sequence of the nematode C. elegans: a platform for  
investigating biology. The C. elegans Sequencing Consortium.";  
RL Science 282:2012-2018 (1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA STRAIN=Bristol N2;  
RA Pauley A., Gattung S.;  
RT "The sequence of C. elegans cosmid F54D11.";



Query Match 16.4%; Score 413; DB 5; Length 475;  
 Best Local Similarity 28.6%; Pred. No. 7.3e-24;  
 Matches 128; Conservative 70; Mismatches 202; Indels 48; Gaps 14;

QY 9 AMLDSDQASDLKKEPEILSMPLPCKGKLELGGAGIGFTGELAKAGAGVIALDFIBS 68  
 DB 24 SMNLNHAEELESDDADILASLPLLNKDDVVDIGAGIGFTTVAETARWVLTDFIDS 83

QY 69 AIKQNEVINGHYKQVPMCADVTSPTLSPPHSLDVFISNMLMYLSDEVENLVERMLX 128  
 DB 84 FIKKQERNNAHLNINQVGDVAG--LKMSNSVDLVFTNMLMYLSDEETVEIFNCR 141

QY 129 WLKPGGYIFFRESCFQHS--GDHKKRS-----NPHYREPRFYTKAFKCHLQDGSNS 180  
 DB 142 WLRSHGIVHLRESCEPSTGRSKAKSMHDANANPHYRFSLSYINLIRAYRDVDNKL 201

QY 181 YELSLSCCKIGAYVRNKNQNOISLWQKVDKDKGFQFLDTSQYKNSILRYERVF 240  
 DB 202 WRFNVQWSCSVPTIYKSNWQVHWAELKVAEDGAGKTSFNEVELIKNTWQNEQEA 261

QY 241 GPGYVSTGGYETTKFVSMMLDKPGQ-----KVLVGGCIGGGDFYMAETDVEVVG 292  
 DB 262 DAKLDEKYYVTDKVFSSALTSLSNSTFLYTPRTVSPYCHINA--HTLAETFNANVMN 319

QY 293 FDLNWMISFALERSIGLK-CAVEF-----EVAD-CTKINYPDNPSPDVIYSRDTLHIOQK 346  
 DB 320 TEIIPYRYTSLTKSNLKDQVRFGWQSLDTSVTYMQKDALFDVFAVTEFLSTVDDE 379

QY 347 PALFRSFTYKWKPGKVLISDYCKKAGPPSPFAAYIKQGYDLHDVKEYGOMLKDAGFV 406  
 DB 380 --TIRQLPNVMSDGAKEPITL-----PVDEVNEAKMKOR-----IQBLGYTLK--SPT 423

QY 407 DVLAE--DRTEQPIRVLRKELETVEKEK 432  
 DB 424 DVTDOCEIAQRYE----KDHEQLRDEK 447

RESULT 11  
 Q23552 ID Q23552 PRELIMINARY; PRT; 475 AA.  
 AC Q23552;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein ZK622.3.  
 GN ZK622.3.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA Wilson R.;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 RL investigating biology. The C. elegans Sequencing Consortium.";  
 EN Science 282:2012-2018(1998).  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Leimbach D.;  
 RT "The sequence of C. elegans cosmid ZK622.";  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Waterston R.;  
 RL EMBL; U39998; AAC038583.1; -  
 DR F1R; T27936; T27936.  
 DR WormPep; ZK622.3a; CE33504.  
 DR GO; GO:0008757; F1S-adenosylmethionine-dependent methyltransf. . . ; IEA.  
 DR InterPro; IPR001609; Methyltransf. . . . .  
 DR InterPro; IPR000051; SAM\_bind.  
 KW Hypothetical protein.  
 SQ SEQUENCE 475 AA; 55147 MW; 0B30BC89651526D9 CRC64;

QY 9 AMLDSDQASDLKKEPEILSMPLPCKGKLELGGAGIGFTGELAKAGAGVIALDFIBS 68  
 DB 39 SMNLNHAEELESDDADILASLPLLNKDDVVDIGAGIGFTTVAETARWVLTDFIDS 98

QY 69 AIKQNEVINGHYKQVPMCADVTSPTLSPPHSLDVFISNMLMYLSDEVENLVERMLX 128  
 DB 99 FIKKQERNNAHLNINQVGDVAG--LKMSNSVDLVFTNMLMYLSDEETVEIFNCR 156

QY 129 WLKPGGYIFFRESCFQHS--GDHKKRS-----NPHYREPRFYTKAFKCHLQDGSNS 180  
 DB 157 WLRSHGIVHLRESCEPSTGRSKAKSMHDANANPHYRFSLSYINLIRAYRDVDNKL 216

QY 181 YELSLSCCKIGAYVRNKNQNOISLWQKVDKDKGFQFLDTSQYKNSILRYERVF 240  
 DB 217 WRFNVQWSCSVPTIYKSNWQVHWAELKVAEDGAGKTSFNEVELIKNTWQNEQEA 276

QY 241 GPGYVSTGGYETTKFVSMMLDKPGQ-----KVLVGGCIGGGDFYMAETDVEVVG 292  
 DB 277 DAKLDEKYYVTDKVFSSALTSLSNSTFLYTPRTVSPYCHINA--HTLAETFNANVMN 334

QY 293 FDLNWMISFALERSIGLK-CAVEF-----EVAD-CTKINYPDNPSPDVIYSRDTLHIOQK 346  
 DB 335 TEIIPYRYTSLTKSNLKDQVRFGWQSLDTSVTYMQKDALFDVFAVTEFLSTVDDE 394

QY 347 PALFRSFTYKWKPGKVLISDYCKKAGPPSPFAAYIKQGYDLHDVKEYGOMLKDAGFV 406  
 DB 395 --TIRQLPNVMSDGAKEPITL-----PVDEVNEAKMKOR-----IQBLGYTLK--SPT 438

QY 407 DVLAE--DRTEQPIRVLRKELETVEKEK 432  
 DB 439 DVTDOCEIAQRYE----KDHEQLRDEK 462

RESULT 12  
 Q23552 ID Q23552 PRELIMINARY; PRT; 484 AA.  
 AC Q23552;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 GN ZK622.3.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA None;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 RL investigating biology. The C. elegans Sequencing Consortium.";  
 EN Science 282:2012-2018(1998).  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Leimbach D.;  
 RT "The sequence of C. elegans cosmid ZK622.";  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Waterston R.;  
 RL EMBL; U39998; AAC038583.1; -  
 DR F1R; T27936; T27936.  
 DR WormPep; ZK622.3a; CE33504.  
 DR GO; GO:0008757; F1S-adenosylmethionine-dependent methyltransf. . . ; IEA.  
 DR InterPro; IPR001609; Methyltransf. . . . .  
 DR InterPro; IPR000051; SAM\_bind.  
 KW Hypothetical protein.  
 SQ SEQUENCE 475 AA; 55147 MW; 0B30BC89651526D9 CRC64;







QY 100 HSLDVFNSNMLVLSDEEVENLVRMLKQKPGGVIFPESCFHQSGDHKRS----- 153  
DB 4 NSVDLVFNMLVLSDEEVEIFPNCNMLASHGIVHRESCEPSTGRSKAKSMHDTA 63  
QY 154 --NPTHVREPRFTYKAFKCHLDQSGNBYELSLSCKICGAYVKNKQNOISLWAKV 211  
DB 64 NANPTHVRFSSLYINLRATRYDNDVKNLWRFNVQMSCVPTIYIKRSNNWQVHLAEKV 123  
QY 212 DSKDDKGFQRFLDTSQYKNSILRYKRVFGVSTGGVYETTKBFPVSMLDLKPQ----- 266  
DB 124 PAEDCAGKTSFNLVELIKNTWQNEAWDAKLDEKYVMTDKVFSALTSLSNSTPFL 183  
QY 267 ---KVLVGGCIGGSDPMAETEDVEVGFDSVNMISFALERSIGLK-CAVEP-----EV 318  
DB 184 YPRVSVSYCHINA--HTLAETFNANWTEIIPYRTSLTKSNLKDQVRFGWQSL 241  
QY 319 AD-CTKINYPNSDVIYSDRTILHIQDKPALFRSFKVWLKPGGVLSIDYCKAGPSP 377  
DB 242 TDSVTYQKQKALPDVFVATEPLSTVDE--TIRQLPNVMSDGAFTILE-----PVDE 293  
QY 378 EFAAYIKQGYDLHDVKEYGQMLKADGFDVVLAE--DRTEQPIRVLRKELETVEKER 432  
DB 294 VNEAEKQR-----IQSLGYTLK--SFTDVTQOCIEAQGVF-----KHQELRDEK 338

RESULT 15  
Q83X76 PRELIMINARY; PRT; 270 AA.  
AC Q83X76;  
DT 01-JUN-2003 (TREMELrel. 24, Created)  
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)  
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
DE Putative NBP-hexose 3-O-methyltransferase.  
OS Streptomyces rochei (Streptomyces parvullus).  
OG Plasmid pSLA2-L.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Streptomyces; Streptomycetaceae; Streptomyces.  
ON NCBI\_TaxID=1928;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=7434AN4;  
RA Mochizuki S., Hiratsu K., Suwa M., Ishii T., Sugino F., Yamada K.,  
RA Kinashi H.;  
RT "The large linear plasmid pSLA2-L of Streptomyces rochei has an  
RT unusually condensed gene organization for secondary metabolism.";  
RL Mol. Microbiol. 0:0-0(2003).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=7434AN4;  
RX MEDLINE=20408175; PubMed=10954087;  
RA Hiratsu K., Mochizuki S., Kinashi H.;  
RT "Cloning and analysis of the replication origin and the telomeres of  
RT the large linear plasmid pSLA2-L in Streptomyces rochei.";  
RL Mol. Gen. Genet. 263:1015-1021(2000).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=7434AN4;  
RX MEDLINE=20231737; PubMed=10767533;  
RA Suwa M., Sugino H., Sasaoka A., Mori E., Fujii S., Shinkawa H.,  
RA Nimi O., Kinashi H.;  
RT "Identification of two polyketide synthase gene clusters on the linear  
RT plasmid pSLA2-L in Streptomyces rochei.";  
RL Gene 246:123-131(2000).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=7434AN4;  
RX MEDLINE=99053144; PubMed=9836424;  
RA Kinashi H., Fujii S., Hatanai A., Kurokawa T., Shinkawa H.;  
RT "Physical mapping of the linear plasmid pSLA2-L and localization of  
RT the eryA and actI homologs.";  
RL Biosci. Biotechnol. Biochem. 62:1892-1897(1998).  
DR EMBL; AB088224; BAC76486.1;  
GO; GO:0046821; C:extrachromosomal DNA; IEA.

DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransferase; IEA.  
DR GO; GO:0016740; F:transferase activity; IEA.  
DR InterPro; IPR001601; Methyltransferase.  
DR InterPro; IPR000051; SAM\_bind.  
DR InterPro; IPR004033; UbiE/COQ5\_Metrf.  
DR Pfam; PF01209; Ubie\_methyltran; 1.  
KW Methyltransferase; Plasmid; Transferase.  
SQ SEQUENCE 270 AA; 29654 MW; F602ECAB8FA82055 CRC64;  
Query Match 9.1%; Score 229.5; DB 2; Length 270;  
Best Local Similarity 29.4%; Pred. NO. 8e-10; 83; Indels 23; Gaps 6;  
Matches 62; Conservative 43; Mismatches 43;  
QY 251 ETTKEFVSMLDLKPQKQKLVGCGIGGDFYMAETFDVEVGFDSVNMISFALERS--I 308  
DB 55 QLTQDLARLAPAPGORILDVGGVGEPAFLARTADVWGVSTISAYQVGRAGERARDF 114  
QY 309 GLKCAVEFEVADCTKINYPNSDVIYSDRTILHIQDKPALFRSFKVWLKPGGVLSIDY 368  
DB 115 GLADRVSFPRHADAELPFPDASTDGGWAFESLIHMPDKEKVLREIKRVLRFGATLVADM 174  
QY 369 CKKAGPSPPEPAAYIKQGYDLHDVKEYGQMLKADG-----FVDVLAEDRTQPIRVLRK 423  
DB 175 FSQ-----PDTELT--QDIITPEMADYRAVIESAGLVVRREFTDITANTLAPPAYR----- 224  
QY 424 ELETVEKE---KDVFIISDFSEEDYNDIVGG 450  
DB 225 --ESVAADLLARKDDVIAMTGPPEFARWVDG 253

Search completed: July 26, 2004, 13:15:27  
Job time : 44 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 2, 2004, 02:21:49 ; Search time 630 Seconds  
(without alignments)

3681.246 Million cell updates/sec

Title:

US-10-031-331B-40

Perfect score:

2515

Sequence:

1 HTVDLITRAMMLDSQASDLD.....KLRTAKGQRWGLFVAKK 473

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3222919 seqs, 2451570024 residues

Total number of hits satisfying chosen parameters: 6445838

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+pin.model -DBV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US10031331/runat\_26072004\_120354\_15090/app\_query.fasta\_1.647  
-DB=Published Applications NA -QFMT=fastcap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blorun62  
-TRANS=human40.cdi -LIST=45 -LOCAL=200 -THR SCORE=pcr -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US10031331 -CGCN 1 1 221 -runat\_26072004\_120354\_15090  
-NCPG=6 -ICPU=3 -NO WAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-XGAPOP=6 -FOAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/PTUS\_PUBCOMB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No. Query Match Length DB ID Description

1	2011	80.0	1884	13	US-10-425-114-14370	Sequence 14370, A
2	1394	78.9	1890	13	US-10-425-114-14613	Sequence 14613, A
3	1932.5	76.8	1911	17	US-10-437-963-53228	Sequence 53228, A
4	1829.5	72.7	1962	15	US-10-149-759-81	Sequence 81, Appl
5	1827	72.6	1897	13	US-10-425-114-24087	Sequence 24087, A
6	1734	68.9	2543	13	US-10-424-599-29759	Sequence 29759, A
7	1321	52.5	1201	13	US-10-425-114-36489	Sequence 36489, A
8	1088	43.3	1267	13	US-10-425-114-2412	Sequence 2412, Ap
9	937	37.3	825	13	US-10-425-114-25382	Sequence 25382, A
10	839	33.4	599	17	US-10-021-323-14225	Sequence 14225, A
11	839	33.4	906	13	US-10-424-599-114653	Sequence 114653, A
12	837	31.7	601	17	US-10-021-323-8285	Sequence 8285, Ap
13	798.5	31.7	586	17	US-10-021-323-13961	Sequence 13961, A
14	768	30.6	694	17	US-10-021-323-17175	Sequence 17175, A
15	768	30.5	525	17	US-10-021-323-14357	Sequence 14357, A
16	719	28.6	487	10	US-09-770-961-776	Sequence 776, App
17	690.5	27.5	1087	17	US-10-437-963-7552	Sequence 7552, Ap
18	687	27.3	500	17	US-10-021-323-17176	Sequence 17176, A
19	670.5	26.7	557	13	US-10-424-599-97831	Sequence 97831, A
20	658	26.2	619	15	US-10-149-759-49	Sequence 49, Appl
21	649	25.8	467	12	US-09-732-627A-1204	Sequence 3204, Ap
22	498	19.8	1416	17	US-10-602-268-18	Sequence 18, Appl
23	498	19.8	1534	17	US-10-602-268-6	Sequence 6, Appl
24	484	19.2	494	10	US-09-770-961-584	Sequence 584, App
25	480	19.1	1311	17	US-10-602-268-17	Sequence 17, Appl
26	480	19.1	1533	17	US-10-602-268-5	Sequence 5, Appl
27	428	17.0	272	9	US-09-294-093B-1679	Sequence 1679, Ap
28	427	17.0	526	13	US-10-424-599-81253	Sequence 81253, A
29	422	16.8	539	17	US-10-021-323-17177	Sequence 17177, A
30	421	16.7	685	17	US-10-437-963-7551	Sequence 7551, Ap
31	420	16.7	290	9	US-09-294-093B-5022	Sequence 5022, Ap
32	405	16.1	591	17	US-10-021-323-3984	Sequence 3984, Ap
33	404	16.1	1380	17	US-10-602-268-13	Sequence 13, Appl
34	402	16.0	1786	17	US-10-602-268-1	Sequence 1, Appl
35	399	15.9	1380	17	US-10-021-323-6613	Sequence 6613, Ap
36	399	15.9	1669	17	US-10-602-268-14	Sequence 14, Appl
37	399	15.7	1371	17	US-10-602-268-2	Sequence 2, Appl
38	396	15.7	1371	17	US-10-602-268-15	Sequence 15, Appl
39	396	15.7	1472	17	US-10-602-268-3	Sequence 3, Appl
40	390	15.5	293	9	US-09-294-093B-160	Sequence 160, App
41	385	15.3	257	9	US-09-294-093B-4860	Sequence 4860, Ap
42	383	15.2	257	9	US-09-923-876-419	Sequence 419, App
43	383	15.2	257	11	US-09-923-876-419	Sequence 419, App
44	375	14.9	530	17	US-10-437-963-53196	Sequence 53196, A
45	353.5	14.1	1407	17	US-10-602-268-16	Sequence 16, Appl

## ALIGNMENTS

## RESULT 1

US-10-425-114-14370

Sequence 14370, Application US/10425114

Publication No. US20040034888A1

GENERAL INFORMATION:

APPLICANT: Liu, Jingdong

APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David K.

APPLICANT: Screen, Steven E.

APPLICANT: Tabaska, Jack E.

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(5333)B

CURRENT APPLICATION NUMBER: US/10/425,114

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 73128

SEQ ID NO 14370

LENGTH: 1884

TYPE: DNA

ORGANISM: Arabidopsis thaliana

FEATURE:

OTHER INFORMATION: Clone ID: LIB22-074-F3\_FLI

US-10-425-114-14370									
Alignment Scores:									
Pred. No.:	5,276-242	Length:	1884						
Score:	2011.00	Matches:	367						
Percent Similarity:	88.56%	Conservative:	51						
Best Local Similarity:	77.75%	Mismatches:	54						
Query Match:	79.96%	Indels:	0						
DB:	13	Gaps:	0						
US-10-031-331B-40 (1-473) x US-10-425-114-14370 (1-1884)									
Qy	1	HiethrValaspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp	20						
Db	256	CAITCCGCTGATCTGATCTGTTGAGCTATGATCTGCTGATCGAGAGCTTCGATCTGCAC	315						
Qy	21	LygluGluArgProGluIleLeuSerMetLeuProProLeuGluGlyLysCysLeuLeu	40						
Db	316	AAGGAGAGAGCTCTGAGGTACTCTCTTTGCTCCCTCATATGAGGCAATATGATGTTG	375						
Qy	41	GluLeuGlyAlaGlyIleGlyArgPheThrGlyGluLeuAlaGluIleGlyGlnVal	60						
Db	376	GAATCTGGAGCTGCTATGCTGCTTTCACTGGTGAATAGCTCAAAAGGCTGGTGAATC	435						
Qy	61	IleAlaLeuAspPheIleGluSerAlaIleLysLysAsnGluValIleAsnGlyHisTyr	80						
Db	436	ATTGCTCTTGACTTCATTTGATACCTTATCAGAGATGATCAAGTATCATGCGGCATAC	495						
Qy	81	LyAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProHis	100						
Db	496	AAGATGTCAGTTTATGCTGCTGATGTTACATCCCTGACCTCAAGATCATGATGGA	555						
Qy	101	SerLeuAspValIlePheSerAsnTrpLeuLeuMetTyrLeuSerAspGluValGlu	120						
Db	556	TCTCTTGACTGATTTCTCACTGCTGCTCATGATCTCTCTCTGACAAAGAGTGAG	615						
Qy	121	AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyGlyTyrIlePheArgGlu	140						
Db	616	CTTTTGGCAGAAAGATGCTGCTGCTGATCAAGTTGGAGGATACATTTCTTCCTGAA	675						
Qy	141	SerCysPheHisGlnSerGlyAspHisLeuArgLysSerAsnProThrHisTyrArgGlu	160						
Db	676	TCTTGTCTCCACCAATCAGGGGACAGTAAAGGGGAAATCCAAACCCCACTCACTACCGTGA	735						
Qy	161	ProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSer	180						
Db	736	CCCCGTTTCTATTCAGAGTCTTTCAAGATGTCAGACTCGGATGCTCTGGAATTC	795						
Qy	181	TyrGluLeuSerLeuLeuSerCysLysCysIleGlyAlaTyrValArgAsnLysLysAsn	200						
Db	796	TTTGAGCTCTATGATCGGATGCAAGTGCAATTGGAGCTTATGTCAGAAACAGAGAGAT	855						
Qy	201	GlnAsnGlnIleSerTrpLeuTrpGlnLysValAspSerLysAspLysLysGlyPheGln	220						
Db	856	CAGAAATCAGATTTGATGATGAGAGAGGTGAGTCAAGAAATCAGAGAGCTTCCAA	915						
Qy	221	ArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArgValPhe	240						
Db	916	CGTTTCTTGACCAATGTCTCAATAATCCAGTGGAAATCCAGTGGAAATCCAGTGGTCTTT	975						
Qy	241	GlyProGlyTyrValSerThrGlyGlyTyrGluThrLysGluPheValSerMetLeu	260						
Db	976	GGCCAAAGGTTTGTGAGCACTGCTGGAATTCAGACCAACCAAGAAATTTGTGGAGAAATG	1035						
Qy	261	AspLeuLysProGlyGlnLysValLeuAspValGlyCysGlyIleGlyGlyLysPhe	280						
Db	1036	ATCTGAAACCAAGCAAGAAATCTAGATGTGGTGTGGCAATTTGGTGGAGGTGATCTC	1095						
Qy	281	TyrMetAlaGluThrPheAspValGluValValGlyPheAspLeuSerValAsnMetIle	300						
Db	1096	TACATGCTGAGAGTTTGTATGTTTACGTTGTGGTATCGATCTTTCTGTCACCATGATC	1155						
Qy	301	SerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluValAlaAsp	320						

US-10-425-114-14370									
Alignment Scores:									
Pred. No.:	5,276-242	Length:	1884						
Score:	2011.00	Matches:	367						
Percent Similarity:	88.56%	Conservative:	51						
Best Local Similarity:	77.75%	Mismatches:	54						
Query Match:	79.96%	Indels:	0						
DB:	13	Gaps:	0						
US-10-031-331B-40 (1-473) x US-10-425-114-14370 (1-1884)									
Qy	1	HiethrValaspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp	20						
Db	256	CAITCCGCTGATCTGATCTGTTGAGCTATGATCTGCTGATCGAGAGCTTCGATCTGCAC	315						
Qy	21	LygluGluArgProGluIleLeuSerMetLeuProProLeuGluGlyLysCysLeuLeu	40						
Db	316	AAGGAGAGAGCTCTGAGGTACTCTCTTTGCTCCCTCATATGAGGCAATATGATGTTG	375						
Qy	41	GluLeuGlyAlaGlyIleGlyArgPheThrGlyGluLeuAlaGluIleGlyGlnVal	60						
Db	376	GAATCTGGAGCTGCTATGCTGCTTTCACTGGTGAATAGCTCAAAAGGCTGGTGAATC	435						
Qy	61	IleAlaLeuAspPheIleGluSerAlaIleLysLysAsnGluValIleAsnGlyHisTyr	80						
Db	436	ATTGCTCTTGACTTCATTTGATACCTTATCAGAGATGATCAAGTATCATGCGGCATAC	495						
Qy	81	LyAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProHis	100						
Db	496	AAGATGTCAGTTTATGCTGCTGATGTTACATCCCTGACCTCAAGATCATGATGGA	555						
Qy	101	SerLeuAspValIlePheSerAsnTrpLeuLeuMetTyrLeuSerAspGluValGlu	120						
Db	556	TCTCTTGACTGATTTCTCACTGCTGCTCATGATCTCTCTCTGACAAAGAGTGAG	615						
Qy	121	AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyGlyTyrIlePheArgGlu	140						
Db	616	CTTTTGGCAGAAAGATGCTGCTGCTGATCAAGTTGGAGGATACATTTCTTCCTGAA	675						
Qy	141	SerCysPheHisGlnSerGlyAspHisLeuArgLysSerAsnProThrHisTyrArgGlu	160						
Db	676	TCTTGTCTCCACCAATCAGGGGACAGTAAAGGGGAAATCCAAACCCCACTCACTACCGTGA	735						
Qy	161	ProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSer	180						
Db	736	CCCCGTTTCTATTCAGAGTCTTTCAAGATGTCAGACTCGGATGCTCTGGAATTC	795						
Qy	181	TyrGluLeuSerLeuLeuSerCysLysCysIleGlyAlaTyrValArgAsnLysLysAsn	200						
Db	796	TTTGAGCTCTATGATCGGATGCAAGTGCAATTGGAGCTTATGTCAGAAACAGAGAGAT	855						
Qy	201	GlnAsnGlnIleSerTrpLeuTrpGlnLysValAspSerLysAspLysLysGlyPheGln	220						
Db	856	CAGAAATCAGATTTGATGATGAGAGAGGTGAGTCAAGAAATCAGAGAGCTTCCAA	915						
Qy	221	ArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArgValPhe	240						
Db	916	CGTTTCTTGACCAATGTCTCAATAATCCAGTGGAAATCCAGTGGAAATCCAGTGGTCTTT	975						
Qy	241	GlyProGlyTyrValSerThrGlyGlyTyrGluThrLysGluPheValSerMetLeu	260						
Db	976	GGCCAAAGGTTTGTGAGCACTGCTGGAATTCAGACCAACCAAGAAATTTGTGGAGAAATG	1035						
Qy	261	AspLeuLysProGlyGlnLysValLeuAspValGlyCysGlyIleGlyGlyLysPhe	280						
Db	1036	ATCTGAAACCAAGCAAGAAATCTAGATGTGGTGTGGCAATTTGGTGGAGGTGATCTC	1095						
Qy	281	TyrMetAlaGluThrPheAspValGluValValGlyPheAspLeuSerValAsnMetIle	300						
Db	1096	TACATGCTGAGAGTTTGTATGTTTACGTTGTGGTATCGATCTTTCTGTCACCATGATC	1155						
Qy	301	SerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluValAlaAsp	320						

Db	1156	TCCTTCGATTCGAAACGCTGCTATTCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGAT	1215
Qy	321	CysThrLysIleAsnTyrProAspAsnSerPheAspValIleTyrSerArgAspThrIle	340
Db	1216	TGCACCACAAAACACTACCCAGATAATTCGTTGATGTCATTTACAGCCCGTGACACTATT	1275
Qy	341	LeuHisIleGlnAspLysProAlaLeuPheArgSerPheTyrLysTrpLeuLysProGly	360
Db	1276	CTGCACATCCAGACAAACGACCTGTTTAGGACTTCTTCAATGGCTTAAACCGGGA	1335
Qy	361	GlyLysValLeuIleSerAspTyrCysLysLysAlaGlyProProSerProGluPheAla	380
Db	1336	GGTAAAGTTCTCATACGCACTACTGTGAAGCCCAAACTCCATCTGCTGAGTTTCA	1395
Qy	381	AlaTyrIleLysGlnArgGlyTyrAspLeuHisAspValLysGlnTyrGlyGlnMetLeu	400
Db	1396	GAGTACATCAACACAGAGAGATATGATCTCCATGACGCTTAAAGCTTATGGAAGATGCTA	1455
Qy	401	LysAspAlaGlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIleArgVal	420
Db	1456	AAAGACCTGCTGCTTCACTGATGTCGACAGGACCGTACTGATCAGTTTATGCAAGTC	1515
Qy	421	LeuArgLysGluLeuGluThrValGluLysGluLysAspValPheIleSerAspPheSer	440
Db	1516	CTGAACGCTGAATTTAGACAGGGTGGAGAAAGAAAGAAATTCCTCCGACTTCTCC	1575
Qy	441	GluGluAspTyrAsnAspIleValGlyGlyTyrAsnAspLysLeuArgThrAlaLys	460
Db	1576	AAAGAGGATTACGATGACATTTGTCAGCATGCAAGTCAAGCTGAGAGGTGTCATCG	1635
Qy	461	GlyGluGlnArgTrpGlyLeuPheValAlaLysLys	472
Db	1636	GATGAGCAGAAATGGGACATTTTCATCCGCCACAG	1671
RESULT 2			
US-10-425-114-14613			
; Sequence 14613, Application US/10425114			
; Publication No. US2004003488A1			
; GENERAL INFORMATION:			
; APPLICANT: Liu, Jingdong			
; APPLICANT: Zhou, Yihua			
; APPLICANT: Kovalic, David K.			
; APPLICANT: Screen, Steven E.			
; APPLICANT: Tabaska, Jack B.			
; APPLICANT: Cao, Yongwei			
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With			
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement			
; FILE REFERENCE: 38-21(53313)B			
; CURRENT APPLICATION NUMBER: US/10/425,114			
; CURRENT FILING DATE: 2003-04-28			
; NUMBER OF SEQ ID NOS: 73128			
; SEQ ID NO 14613			
; LENGTH: 1890			
; TYPE: DNA			
; ORGANISM: Arabidopsis thaliana			
; FEATURE:			
; OTHER INFORMATION: Clone ID: LIB23-061-D6_FLI			
US-10-425-114-14613			
Alignment Scores:			
Pred. No.:	1,328-238	Length:	1890
Score:	1384.00	Matches:	366
Percent Similarity:	86.35%	Conservative:	51
Best Local Similarity:	77.54%	Mismatches:	55
Query Match:	78.89%	Indels:	1
DB:	13	Gaps:	0
US-10-031-331B-40 (1-473) x US-10-425-114-14613 (1-1890)			
Qy	1	HiethrValaspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp	20
Db	249	CAITCCGCTGATCTGATCTGTTGAGCTATGATCTGCTGATCGAGAGCTTCGATCTGCAC	308

```

QY 391 AlaTyrIleLysGlnArgGlyTyrAspLeuHisAspValLysGluTyrGlyGlnMetLeu 400
Db 1388 GAGTATCATCAAAACAGAGAGGATATGATCTCATGAGGTTCAAGACTTATGGACAGATGCTA 1447
QY 401 LysAspAlaGlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIleArgVal 420
Db 1448 AAGAAGCGCTGGCTTCACGTATGTCATCGCAGAGACCGTACTGATCAGTTTATCGAAGTC 1507
QY 421 LeuArgLysGluLeuGluThrValGluLysGluLysAspValPheIleSerAspPheSer 440
Db 1508 CTCGAACGCTGANTTAGACAGCGTGGAGAAAGAAAGAAAGAAATTCATCTCCGACTTCTCC 1567
QY 441 GluGluAspTyrAsnAspIleValGlyGlyTyrPAsnAspLysLeuArgArgThrAlaLys 460
Db 1568 AAGAAGGATTACGATGACATTTGTTGGAGATGCAAGTCAAAGCTGAGAGGATGTCATCG 1627
QY 461 GlyGluGlnArgTyrGlyLeuPheValAlaLysLys 472
Db 1628 GATGACGAGMAATGGGCACTTTTCATCGCCCAACAG 1653

RESULT 3
US-10-437-963-53228
; Sequence 53228, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated
; WITH THE INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 53228
; LENGTH: 1911
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_5544C.1
US-10-437-963-53228

Alignment Scores:
Pred. No.: 4,08e-232 Length: 1911
Score: 1932.50 Matches: 359
Percent Similarity: 81.71% Conservative: 52
Best Local Similarity: 71.37% Mismatches: 61
Query Match: 76.84% Indels: 31
DB: 17 Gaps: 2

US-10-031-331B-40 (1-473) x US-10-437-963-53228 (1-1911)

QY 1 HisThrValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp 20
Db 82 CACTCCAAGGACCTTCACCGCTCGAGGGCATTGATGCTCGACCTCCGCGCGCGGATCTCGAC 141
QY 21 LysGluGluArgProGlu 26
Db 142 AAGAGGAGGCGCCCGGAGATCTGGGGGCGCGTGATGCACTTTTGTGTACTTTTCCCAACC 201
QY 27 -----IleLeuSerMetLeuProLeuGluGlyLysCysLeu 39
Db 202 AAATACAGATTCTAGCTGCTAATATTTGTTTACTTCTCTTACGAGGAATATCAGTA 261
QY 40 LeuGluLeuGlyValAspIleGlyArgPheThrGlyLeuAlaGluLysAlaGlyCln 59
Db 262 CTGGAACTTGGTGGTAATAGTGGCTTCACTGSGAATCTAGTGAACAACAGCTGGGGCAT 321

```



655 AACACATCGATTTCAATGTGGGATGTGACCTCTCCAGACCTGATATTTGACGAGGT 714  
 QY 101 SerLeuAspValIlePheSerAsnTrpLeuMetTyrLeuSerAspGluValGlu 120  
 Db 715 TCTGGGATCTGGTGTCTCAATTTGGCTTCTCATGTACTTGTCTGACGAGAGGTAAA 774  
 QY 121 AsnLeuValGluArgMetLeuTyrTrpLeuTyrProGlyTyrIlePheAsnGlu 140  
 Db 775 GGTCTAGCATCAGCGGTATGATGCTGAGCTGAGCTGAGGATACATTTTCTCAGAGAA 834  
 QY 141 SerCysPheHisGlnSerGlyPheHisGlySerAsnProThrHisTyrArgGlu 160  
 Db 835 TCTCTCTTCCACCTAGTACGAGATCAACAGCAGAGCAATCTCTACTACCTACCTCAA 894  
 QY 161 ProArgPheTyrTrpLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSer 180  
 Db 895 CCAACAGAGTACAGCAATCTTCCAGCAGCCTACATCGAGAG---GATGGTCTCTAT 951  
 QY 181 TyrGluLeuSerLeuSerCysLysCysIleGlyAlaTyrValArgAsnLysLysAsn 200  
 Db 952 TTCAGGTTGAANTGGTGGATGCAATGTGCGCACATACGTGCGAAATGAAGAAAT 1011  
 QY 201 GlnAsnGlnIleSerTrpLeuTrpGlnLysVal-----AspSerLysAspLysGly 218  
 Db 1012 CAAACCCAGGTGTGTGTGTATGAGGAAAGTTTCAGTGGATGCGACCTGACGAGCGAGTGT 1071  
 QY 219 PheGlnArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArg 238  
 Db 1072 TTCAGAGATTTTGGACACCCACAGTACACGTCACTGGATCTCTGCTTACGAGCGT 1131  
 QY 239 ValPheGlyProGlyTyrValSerThrGlyTyrGluThrThrLysGluPheValSer 258  
 Db 1132 ATTCTTGGAGAGAGTTTGTAGCAGGCTGGATCGAAACACCAAGAGCTTTGTAGT 1191  
 QY 259 MetLeuAspLeuLysProGlyGlnLysValLeuAspValCysGlyIleGlyGly 278  
 Db 1192 ATGCTGAGTTGAAGCCAGCAGCAGCTGCTTGTGAGCTTGGATGGGATCGAGGTGT 1251  
 QY 279 AspPheTyrMetAlaGluThrPheAspValGluValValGlyPheAspLeuSerValAsn 298  
 Db 1252 GATTTCTACATGCGCGAGAGATATGATGTGAAGTTGTGGCATCGACTGTCTTAAT 1311  
 QY 299 MetIleSerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluVal 318  
 Db 1312 ATGATTTCTGTGCTTGTGAACATCGATCGGACAGAAATGTGCACTGAGTGTGAAGT 1371  
 QY 319 AlaAspCysThrLysIleAsnTyrProAspAsnSerPheAspValIleTyrSerArgAsp 338  
 Db 1372 GGGGATTCACCAAGATTAATACCTCAGCGATCTTTTGTATGATCACTACAGTCGTGAT 1431  
 QY 339 ThrIleLeuHisIleGlnAspLysProAlaLeuPheArgSerPheTyrLysTrpLeuLys 358  
 Db 1432 ACCATTCACATTCAGATTAACCTGCGCTTTTTCACACGTTTATTAATGTTGAG 1491  
 QY 359 ProGlyLysValLeuIleSerAspTyrCysLysLysAlaGlyProTofSerProGlu 378  
 Db 1492 CCTGGAGTCCGGTGTATCAGTCACTAGTCACTGCTAGAGCTCCACAACTCGTCCGCGAG 1551  
 QY 379 PheAlaAlaTyrIleLysGlnArgGlyTyrAspLeuHisAspValLysGluTyrGln 398  
 Db 1552 TTGCTGCATACATTCAGACAGGGGTATGATCTCCATACCGTTTATTAATGTTGAG 1611  
 QY 399 MetLeuLysAspAlaGlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIle 418  
 Db 1612 ATGCTGAAGATCCCGTTTGTGAAGTGTGCGAGAGGACCGCAGCGATCAGTTTCA 1671  
 QY 419 ArgValLeuArgLysGluLeuGluThrValGluLysGluLysAspValPheIleSerAsp 438  
 Db 1672 GAAGTGTTCAGAGGAGTACCCACCTAGACGAGTGTGACCGAGTTCATCACAGAT 1731  
 QY 439 PheSerGluLysThrAsnAspIleValGlyTyrAsnAspLysLeuArgThr 458  
 Db 1732 TTCTCCAGAGAGATTATACTACATTTGTGAGCGGATGAGAGTAACTGAGCGCTGT 1791

QY 459 AlaLysGlyGluGlnArgTrpGlyLeuPheValAlaLysLys 472  
 Db 1792 TCGAATGACGACAGAGTGGGACTCTTTCATAGCTACGAC 1833

## RESULT 5

US-10-425-114-24087  
 ; Sequence 24087, Application US/10425114  
 ; Publication No. US2004003488A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jiqdong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovacic, David K.  
 ; APPLICANT: Screen, Steven E.  
 ; APPLICANT: Tabaska, Jack E.  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(5313)B  
 ; CURRENT APPLICATION NUMBER: US/10/425.114  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 73128  
 ; SEQ ID NO 24087  
 ; LENGTH: 1897  
 ; TYPE: DNA  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: LIB3607-019-D6\_FLI  
 US-10-425-114-24087

Alignment Scores:  
 Pred. No.: 7,62e-219 Length: 1897  
 Score: 1827.00 Matches: 341  
 Percent Similarity: 83.47% Conservative: 53  
 Best Local Similarity: 72.25% Mismatches: 78  
 Query Match: 72.64% Indels: 1  
 DB: 13 Gaps: 0

US-10-031-331B-40 (1-473) x US-10-425-114-24087 (1-1897)

QY 1 HisThrValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp 20  
 Db 240 CACTCCGGGAGCTCAACCTCGAGGCGCATATGCTGACTCCGCGCCGCGCAACTCGAC 299  
 QY 21 LysGluGluArgProGluIleLeuSerMetLeuProLeuGluGlyLysCysLeuLeu 40  
 Db 300 AAGGAGGAGCGCCCGAGGTTCTGCTTCTTCTCTATATGAGGGAATCTATCTG 359  
 QY 41 GluLeuGlyAlaGlyIleGlyArgPheThrGlyGluLeuAlaGluLysAlaGlyGlnVal 60  
 Db 360 GAGCTGGGAGCTGGAATAGGCGCTTACTGTGTAACCTGGCTTAACATCTGGGCGATGT 419  
 QY 61 IleAlaLeuAspPheIleGluSerAlaIleLysLysAsnGluValIleAsnGlyHisTyr 80  
 Db 420 TTTCGAGTGAATTCGTTTGAAGTGTGATTAAAGATGGAAGTATAAATGATCACTAT 479  
 QY 81 LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProHis 100  
 Db 480 GGCAACACATCTTTATGTGCTGATGTATACATCCCGGACCTGATGATGAGCAAC 539  
 QY 101 SerLeuAspValIlePheSerAsnTrpLeuMetTyrLeuSerAspGluGluValGlu 120  
 Db 540 TCAATTGATCTGATTTTCAAACTGGTGTCTGATGATCTTTTCAGATGAGGAGATTGAC 599  
 QY 121 AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyTyrIlePhePheAsnGlu 140  
 Db 600 AAGTTCGTGAAGAAGATGTTAAATGTTGAAGTCTGCTGTTATCTTCTTTAGGGA 659  
 QY 141 SerCysPheHisGlnSerGlyAspHisLysArgLysSerAsnProThrHisTyrArgGlu 160  
 Db 660 TCTTGCTTCATCANTCCGAGATACAGAAAGGAATTTTATCCACACATCTACGAGAA 719  
 QY 161 ProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSer 180



TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53223)B

CURRENT APPLICATION NUMBER: US/10/424,599

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 285684

SEQ ID NO 29759

LENGTH: 2543

TYPE: DNA

ORGANISM: Glycine max

FEATURE:

OTHER INFORMATION: Clone ID: PAT\_MRT3847\_126875C.1

US-10-424-599-29759

US-10-031-331B-40 (1-473) x US-10-424-599-29759 (1-2543)

Alignment Scores:

Pred. No.: 6,19e-207

Score: 1734.00

Percent Similarity: 82.35%

Best Local Similarity: 67.86%

Query Match: 58.95%

DB: 13

US-10-031-331B-40 (1-473) x US-10-424-599-29759 (1-2543)

QY 1 HisthrValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp 20

DB 593 CATTCGCGCGATCTCTCGTCGAGGCAATGATGCTCGATTCGAAGCGCGCTCATCTCGAC 652

QY 21 LysGluGluArgProGluIleLeuSerMetLeuProLeuGluGlyLysCysLeuLeu 40

DB 653 AAGAAGAGAGACCGAGGCTCTTCCCTATTACCTCCATTCGAGGCAAGTCCGTATT 712

QY 41 GluLeuGlyAlaGlyIleGlyArgPheThrGlyGluLeuAlaGlyLysAlaGlyGlnVal 60

DB 713 GAGCTTGT 772

QY 61 IleAlaLeuAspPheIleGluSerAlaIleLysLysAlaGluValIleAlaGlyHisTyr 80

DB 773 CTTGCTGTGACTTCATTGACATCTGCTATCAAGAGAAATGAACCAATCAATGACACAC 832

QY 81 LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSer-----PhePro 98

DB 833 AACCATGTCAATTTCTGT 892

QY 99 ProHisSerLysAspValIlePheSerAsnTrpLeuMetMetTyrLeuSerAspGluGlu 118

DB 893 GAAGGATCGGT 952

QY 119 ValGluAsnLeuValGluArgMetLeuLysTrpLeuLysProGlyGlyTyrIlePhePhe 138

DB 953 GTTCAAAATTTACTGACAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1012

QY 139 ArgGluSerCysPheHis-GlnSerGlyAspHisLysLysSerAsnProThrHisTyr 158

DB 1013 AAGGAACTGT 1072

QY 158 rArgGluProAspPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySer-G 178

DB 1073 CAAGGTACCCAGATTTAAATACCAAGCAATTAAGAGAGTGCATATGATGATGATGATGAT 1132

QY 178 LysAsnSerTyrGluLeuSerLeuLeuSerCysLysCysLysLysLysLysLysLysLys 198

DB 1133 GAAATTCCTTTGAGCTTTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1192

QY 198 YsLysAsnGlnAsnGlnIleSerTrpLeuTrpGlnLysValAspSerLysAspLysG 218

DB 1193 AGAAGAAATCAAAACCAAGATTTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1252

QY 218 LysPheGlnArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGlu 238

DB 1253 GTTTCAGCGCTCTCTGT 1312

QY 238 rGValPheGlyProGlyTyrValSerThrGlyGlyTyrGluThrLysGluPheValS 258

DB 1312

QY 1312

DB 1312

QY 1312

DB 1312

QY 1312

DB 1312

QY 1312

DB 1312

QY 1312

DB 779

QY 200

DB 839

QY 220

DB 899

QY 240

DB 959

QY 260

DB 1019

QY 280

DB 1078

QY 300

DB 1138

QY 320

DB 1198

QY 340

DB 1258

QY 360

DB 1318

QY 380

DB 1378

QY 400

DB 1438

QY 420

DB 1498

QY 440

DB 1558

QY 460

DB 1618

QY 472

DB 1654

QY 472

DB 1654

QY 472

DB 1654

QY 472

DB 1654

QY 472

DB 1654

QY 472

DB 1654

QY 472

DB 1654

QY 472

DB 1654



Db 1313 GGAATGATGCGCCAGGCTTTGTGAGCACTGTGTGAAACGACAAAGAGATTGTGG 1372  
 Qy 258 erMetLeuaspLeuysProGlyGlnlyValLeuAepValGlyCysGly11leGlyGlyG 278  
 Db 1373 CAAAGCTGGACCTAAAGCTGTGAGAAAGTTCTGGATGTGTGGTGTGTGGGAG 1432  
 Qy 278 lyAspPheTyMetAlaGluThrPheAepValGluValValGlyPheAepLeuSerVala 298  
 Db 1433 GTGATTTCTACATGCTGAGAAATTTGATGTTGAGGTATTGTCATTGACCTCTCATAA 1492  
 Qy 298 smMetileSerPheAlaLeuGluArgSerileGlyLeuysCysAlaValGluPheGluV 318  
 Db 1493 ACATGNTTCTTGTGCTATGAGAGCTGTATGTGACTGAACCTAGCTGTGTGAATTGAT 1552  
 Qy 318 alAlaAepCysThrylileAenTyProAepAenSerPheAepValileTySerArgA 338  
 Db 1553 GTGCGGATGCTATAGAAAACATATCTCTGAGAAATCATATTGATGATATATACATCGG 1612  
 Qy 338 spThrylileLeuHisileGlnAepLysProAlaLeuPheArgSerPheTyTySerArgA 358  
 Db 1613 ACACCATGCTGACGTCAAGATTAACCAACATATTATTCAGATCATTTTACAAGTGTGA 1672  
 Qy 358 yepProGlyGlyValleuileSerAepTyTyCysLysLysAlaGlyProProSerProG 378  
 Db 1673 ACCCTGGAGGAAATTTCTATCAAGATTTACTGCAAAAGTGTGGAAGTCCATCTTAG 1732  
 Qy 378 luPheAlaAlaTyrlileysGlnArgGlyTyAspLeuHisAepVallyGluTyGlyG 398  
 Db 1733 AATTTGCTGAGTACATATAAAGAGGATTAATCTCCATGACATTAAGCGGTATAGC 1792  
 Qy 398 InMetLeuLysAlaGlyPheValAepValleuAlaGluAepArgThryGluInPheI 418  
 Db 1793 AGATCTTGAGATGCGGATTTGATGATGTCATGTCGAGAGTGAACCTGATCATGTTG 1852  
 Qy 418 leArgValLeuArgGlyGluLeuGluThrValGluLysGlyLysAepValPheileSera 438  
 Db 1853 TGAAACAGCTACAGCAGGAGTAAATGCCCTTGAGAAACAGAGGAGCGATTTATTGCTG 1912  
 Qy 438 spPheSerGluGluAepTyAsnAepileValGlyGlyTyPheAepLysLeuArgArg 458  
 Db 1913 ACTTCAGCGAGAGATACATAGATTTGCAAGATGCAAGAGGAGGAGGAGGAGG 1972  
 Qy 458 hrAlaLysGlyGluGluArgGlyLeuPheValAlaLysLys 472  
 Db 1973 GTGCATCTAGAGAGCAGATGCGGCTTGTTCATTGCCAAGAA 2016

## RESULT 7

US-10-425-114-36489  
 / Sequence 36489, Application US/10425114  
 / Publication No. US20040034888A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Liu, Jingdong  
 / APPLICANT: Zhou, Yihua  
 / APPLICANT: Kovalic, David K.  
 / APPLICANT: Screen, Steven E  
 / APPLICANT: Tabaska, Jack E  
 / APPLICANT: Cao, Yongwei  
 / TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
 / FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 / CURRENT APPLICATION NUMBER: US/10/425,114  
 / NUMBER OF SEQ ID NOS: 73128  
 / SEQ ID NO 36489  
 / LENGTH: 1201

TYPE: DNA

ORGANISM: Zea mays subsp. mexicana

FEATURE:

OTHER INFORMATION: Clone ID: UC-ZMOT0505INTE108B03\_FLI

US-10-425-114-36489

Alignment Scores:

Pred. No.:

1.74e-155 Length: 1201

RESULT 8

Score: 1321.00 Matches: 241  
 Percent Similarity: 87.26% Conservative: 33  
 Best Local Similarity: 76.75% Mismatches: 40  
 Query Match: 52.52% Indels: 0  
 DB: 13 Gaps: 0

US-10-031-331B-40 (1-473) x US-10-425-114-36489 (1-1201)

Qy 159 ArgGluProArgPheTyThrylLeAlaPheLysGluCysHisLeuGlnAepGlySerGly 178  
 Db 1 CGAGAACCAAGGTTTATACCAAGGTTTAAAGAGGCGCATTCATTTCATCAGATGA 60  
 Qy 179 AenSerTyGluLeuSerLeuLeuSerCysLysCysValleGlyAlaTyTyValArganLys 198  
 Db 61 GGTTCGTTTGAACCTTCTTAGTGACCTGTAAATGATTTGGGCTTTATGTCACAAACAG 120  
 Qy 199 lysAenGlnAenGlnileSerTyPheLysGlnlyValAaspSerTyAaspLysGly 218  
 Db 121 AAGAAATCAAAACAGATATGCTGTTATGGAAGAGTAAATCAACAGACAGAGAT 180  
 Qy 219 PheGlnArgPheLeuAepThrySerGlnTyTyLysCysAenSerTyLeuAepTyGluArg 238  
 Db 181 TTTCAAGATTCCTGGACCAAGTGCATACAAACAGTGGGATATTACGTTATGAGGT 240  
 Qy 239 ValPheGlyProGlyTyTyValSerThryGlyTyTyGluThryThryGluPheValSer 258  
 Db 241 GTCTTTGTGAGGTTTGTGAGCCTGTGTGGAATCGAGACTACAAAGCAATTTGTGGGC 300  
 Qy 259 MetLeuAepLeuysProGlyGlnLysValLeuAaspValGlyCysGly11leGlyGly 278  
 Db 301 ATGCTGCACTTAAACCGGCGCAGAAAGTACTTGATGTCGATGTGAATTTGGAGCGGC 360  
 Qy 279 AspPheTyTyMetAlaGluThryPheAepValGluValValGlyPheAepLysSerValAen 298  
 Db 361 GACTTTACATGCTGCAAACTATGATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 420  
 Qy 299 MetileSerPheAlaLeuGluArgSerileGlyLysLysCysAlaValGluPheGluVal 318  
 Db 421 ATGCTTTCATTTGCAATTTGAACGTCGCAATTTGAGCGCAAGTGTCTGTTGAATTCGAGTT 480  
 Qy 319 AlaAepCysThrylileAenTyProAepAenSerPheAepValileTyTySerArgA 338  
 Db 481 GCTGATTCGACCAAGAGATTAACCCAGAAATAGTTTACGTCATCTACAGCGTGAC 540  
 Qy 339 ThrileLeuHisileGlnAepLysProAlaLeuPheArgSerPheTyTyTyTyTyTyTyTy 358  
 Db 541 ACCATCTTCATACATACAGACAGGCTGCTGCTGTTTCAAGAGCTTCTTCAATAGCTTAAG 600  
 Qy 359 ProGlyLysValleuileSerAepTyTyCysLysLysAlaGlyProProSerProGlu 378  
 Db 601 CCGCGCGCAAGTCTCTTAATCAGCGACTACTGTGAAGATCCTCGAAACCATCAGAGAA 660  
 Qy 379 PheAlaAlaTyTyLysGlnArgGlyTyAepLeuHisAepVallyGluTyGlyGln 398  
 Db 661 TTTGCTGCTACATTAAAGCAGAGAGGCTATGACCTTCAGAGTGAAGGCTTATGGACAG 720  
 Qy 399 MetLeuysAepAlaGlyPheValAepValLeuAlaGluAepArgThryGluGluPheile 418  
 Db 721 ATGCTGAAGGATGCTGTTTTCATATATGTCATCGCGAAGATGCGACTGACGAGTCTCTG 780  
 Qy 419 ArgValLeuArgLysGluLeuGluThryValGluLysGlyLysAepValPheileSerAep 438  
 Db 781 AATGTTCTACAGAGGAGATAGTGAAGTTTGAAGAAACAAAGAGCGCTTCTCTGCGAGAC 840  
 Qy 439 PheSerGluAepTyTyAenAepileValGlyGlyTyPheAenAepLysLeuArgArgThry 458  
 Db 841 TTCACCGAGAGGACTATGACGACATTTGTGAATGCTGGAACCGGAGCTGAAACCGAGC 900  
 Qy 459 AlalyeGlyGluGlnArgTyGlyLeuPheValAlaLysLys 472  
 Db 901 TCTGCGCGGAGCAGAGGTGGGGTGTGTTTCATTGCCACCAAG 942

US-10-425-114-2412  
; Sequence 2412, Application US/10425114  
; Publication No. US2004003488A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 2412  
; LENGTH: 1267  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700211781\_FLI  
US-10-425-114-2412  
Alignment Scores:  
Pred. No.: 3,98e-126 Length: 1267  
Score: 1088.00 Matches: 210  
Percent Similarity: 76.54% Conservative: 38  
Best Local Similarity: 64.81% Mismatches: 68  
Query Match: 43.28% Indels: 8  
DB: 13 Gaps: 3  
US-10-031-331b-40 (1-473) x US-10-425-114-2412 (1-1267)  
QY 153 SerAspProThrHis-----TyrArgGluProArgPheTyrThrLysAlaPhe 168  
DB 34 TCAAGCCACATACATCTGGTTTGGATCTTCTTGTATGATGTTTACATAAGCTCCATGCTT 93  
QY 169 LysGluCysHisLeuGlnAspCysSerGlyAsnSerTyrGluLeuSerLeuSerCys 188  
DB 94 ---GATGTAATCAAGAAAGTTTGAAGAAAGTCCATC-----CTGATTAAGCT 141  
QY 189 LysCysIleGlyAlaTyrValArgAsnLysLysAsnGlnAsnGlnLeuSerTyrLeuTrp 208  
DB 142 TTGGCATCTTTAAATACCTTGGTATAAACCTTGGTTCTCGATAGTGTGGTCTATGG 201  
QY 209 GlnLysValAspSerLysAspLysGlyPheGlnArgPheLeuAspThrSerGlnTyr 228  
DB 202 AAAAAGTAACATCATCAGACAGATGGGGATTTTCAAAGTTTTTGGCAATGTGCAATAC 261  
QY 229 LysCysAsnSerIleLeuArgTyrGluArgValPheGlyProGlyTyrValSerThrGly 248  
DB 262 AAAGCCACTGGAATACCTAGCTATGAACGTATCTTTGGAGATGGCTACGTGAGTACTGT 321  
QY 249 GlyTyrGluThrThrLysGluPheValSerMetLeuAspLeuLysProGlyGlnLysVal 268  
DB 322 GGAGCTGAGACTACAAAGATTTGTGGGAAACTGATCTTAAGCTCTGGGCAAGAGTG 381  
QY 269 LeuAspValGlyCysGlyIleGlyAspPheTyrMetAlaGluThrPheAspVal 288  
DB 382 CTTCAGCTTGGATGTGGAATTTGGGGAGGTGACTTTTATATGCTGGAAGATGATGTACA 441  
QY 289 GluValValGlyPheAspLeuSerValAsnMetIleSerPheAlaLeuGluArgSerIle 308  
DB 442 CATGTTGTTGGTATTGACCTTTTCCATTAACATGATAAATGTTTGGCTTGGAGCGTTCATT 501  
QY 309 GlyLeuLysCysAlaValGluPheGluValAlaAspCysThrLysIleAsnTyrProAsp 328  
DB 502 GGGTGAAGTCTTAGTGTGAGTTTGAAGTTTGGATGTCACCAACAGACATACCCAGAC 561  
QY 329 AsnSerPheAspValIleTyrSerArgAspThrIleLeuHisIleGlnAspLysProAla 348  
DB 562 CACATGTTTGAATGTCATCTACATGTCGACACTCTCTTCATATACAGATAAACCTCC 621

QY 349 LeuPheArgSerPheTyrLysTrpLeuLysProGlyGlnLysValLeuLeuSerAspTyr 368  
DB 622 TTGTTTAAAGAGTTTCTTCAAAATGGCTGAAACCTGGGGGAAAGGTTCTTAATCAGTGATAC 681  
QY 369 CysLysLysAlaGlyProPheProSerProGluPheAlaLeuLysIleLysGlnArgGlyTyr 388  
DB 682 TGCAAGAGTCTCTGGAACCAATCAGAACAGAGTTTGCACATACATATTAAGCAGAGGGGTAT 741  
QY 389 AspLeuHisAspValLysGluTyrGlyGlnMetLeuLysAspAlaGlyPheValAspVal 408  
DB 742 GATCTCATGATGTGGAGGCTTATGACAGATGCTCAGATGCTGGTTTCAGTCATGTC 801  
QY 409 LeuAlaGluAspArgThrGluGlnPheIleArgValLeuArgLysGluLeuGluThrVal 428  
DB 802 ATTGCTGAAGACCGAAGTCTCCTCAGTGTGTTTACAGAGGAGCTAGACAAATTT 861  
QY 429 GlnLysGluLysAspValPheIleSerAspPheSerGluGluAspTyrAsnAspIleVal 448  
DB 862 CAGAAGAACAAAGATGATTTCTGCTGAGTTTGGCCAGGAGGATTTATGACGATATCGTG 921  
QY 449 GlyGlyTyrAsnAspLysLeuArgArgThrAlaLysGlyGluGlnArgTyrGlyLeuPhe 468  
DB 922 AATGATGTGAAGGCAAACTGACAGAGCTCTGCTGCTGACAGAGGTGGGGGCTGTTC 981  
QY 469 ValAlaLysLys 472  
DB 982 GTGGCCACCAAG 993  
RESULT 9  
US-10-425-114-25382  
; Sequence 25382, Application US/10425114  
; Publication No. US2004003488A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 25382  
; LENGTH: 825  
; TYPE: DNA  
; ORGANISM: Gossypium hirsutum  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3829-031-P4\_FLI  
US-10-425-114-25382  
Alignment Scores:  
Pred. No.: 2,01e-107 Length: 825  
Score: 937.00 Matches: 176  
Percent Similarity: 88.07% Conservative: 16  
Best Local Similarity: 80.73% Mismatches: 26  
Query Match: 37.26% Indels: 0  
DB: 13 Gaps: 0  
US-10-031-331b-40 (1-473) x US-10-425-114-25382 (1-825)  
QY 255 GluPheValSerMetLeuAspLeuLysProGlyGlnLysValLeuAspValGlyCysGly 274  
DB 2 GAATTTTGGCGAAGTTGGATCTAAAGCTCTGCAAGAGGTTCTAGAGTTGGCTGTGGC 61  
QY 275 IleGlyCysIleAspPheTyrMetAlaGluThrPheAspValGluValValGlyPheAsp 294  
DB 62 ATAGGTGAGGTGACTTTTATATGTCAAAGAAATTTGGAGTCTATGTTGTGGGCTCGAT 121  
QY 295 LeuSerValAsnMetIleSerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaVal 314

```

122 CTGTCATAAATATGATATCTTTGCTCTTGAAGAGCCCAATGCACTCGATTGCTGTG 181
315 GluPheGluValAlaAspCysThrLysLeuAsnTyrProAspAsnSerPheAspValIle 334
182 GAATTTGAAGTTGCTGATTCACAAAGACTTATCCAGATACAGCTTTCATGTATC 241
335 TyrSerArgAspThrIleLeuHisTleGlnAspLysProAlaLeuPheArgSerPheTyr 354
242 TACAGCCGTGACAGATCTTACACATTCAGACAGAGCTGACATTTAGATATTCTAC 301
355 LysTrpLeuLysProGlyLysValLeuLysLeuLysSerAspTyrCysLysLysAlaGlyPro 374
302 AAATGTTGAAGCCAGGGGGGCAAGTTCTTATAGTGTATCTACCAAGAGCGCAGCT 361
375 ProSerProGluPheAlaAlaTyrIleLysGlnAspLysValLeuAlaGluAspArgThr 414
362 GCTTATGCGCAGATCTTGAAGGATGCTGTGTTTGTATCAGGTCAATTCAGAGGATCGAAC 481
415 GluGlnPheLeuArgValLeuArgLysGluLeuGluThrValGluLysGluLysAspVal 434
482 GATCAGTTTCATACAAAGTTCTGCAGCGTGAATTAGACAAAGTTGAGAGGACGAGTACC 541
435 PheLeuSerAspPheSerGluGluAspTyrAsnAspIleValGlyTyrPheAspLys 454
542 TTCATCAGCGACTTTCTCAGAGAGACTACAGAGATATCGTTGGGGGATGGAAGGCAAG 601
455 LeuArgArgThrAlaLysGlyGluGlnArgTyrGlyLeuPheValAlaLysLys 472
602 CTGATTAGACCAAGATTTGGGGAGCAGAGATGGGGTCTTTTTCATTGCCAACAA 655

```

## RESULT 10

```

US-10-021-323-14225
: Sequence 14225, Application US/10021323
: Publication No. US20040123340A1
: GENERAL INFORMATION:
: APPLICANT: Deikman, Jill
: APPLICANT: Feng, Paul C.C.
: APPLICANT: Fincher, Karen L.
: APPLICANT: Ziegler, Todd E.
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
: FILE REFERENCE: 38-21(52274)B
: CURRENT APPLICATION NUMBER: US/10/021,323
: PRIOR FILING DATE: 2001-12-12
: PRIOR FILING DATE: 2000-12-14
: NUMBER OF SEQ ID NOS: 17880
: SEQ ID NO 14225
: LENGTH: 599
: TYPE: DNA
: ORGANISM: Gossypium hirsutum
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (1)..(599)
: OTHER INFORMATION: unsure at all n locations
: OTHER INFORMATION: Clone ID: LIB3829-031-Q6-K6-F4
US-10-021-323-14225

```

## Alignment Scores:

```

Pred. No.: 2,58e-95 Length: 599
Score: 839.00 Matches: 161
Percent Similarity: 89.34% Conservative: 15
Best Local Similarity: 91.73% Mismatches: 21
Query Match: 33.36% Indels: 1
DB: 17 Gaps: 0

```

US-10-031-331B-40 (1-473) x US-10-021-323-14225 (1-599)

```

255 GluPheValSerMetLeuAspLeuLysProGlyGlnLysValLeuAspValGlyCysGly 274
3 GAATTTGTCGCGAATTTGATCTTAAAGCTGGCCA-AGGGTCTAGATGTGGCTGTGGC 61
275 IleGlyGlyAspPheTyrMetAlaGluThrPheAspValGluValValGlyPheAsp 294
62 ATAGGTGGAGGTGACTTTTATATGCAAAAGAAATTTGGAGTCTATGTGTGGGCATCGAT 121
295 LeuSerValAsnMetIleSerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaVal 314
122 CTGTCATAAATATGATATCTTTGCTCTTGAAGAGCCCAATGCACTCGATTGCTGTG 181
315 GluPheGluValAlaAspCysThrLysLeuAsnTyrProAspAsnSerPheAspValIle 334
182 GAATTTGAAGTTGCTGATTCACAAAGACTTATCCAGATACAGCTTTCATGTATC 241
335 TyrSerArgAspThrIleLeuHisTleGlnAspLysProAlaLeuPheArgSerPheTyr 354
242 TACAGCCGTGACAGATCTTACACATTCAGACAGAGCTGATGATCTCCATGATGTATA 301
355 LysTrpLeuLysProGlyLysValLeuLysLeuLysSerAspTyrCysLysLysAlaGlyPro 374
302 AAATGTTGAAGCCAGGGGGGCAAGTTCTTATAGTGTATCTACCAAGAGCGCAGCT 361
375 ProSerProGluPheAlaAlaTyrIleLysGlnArgTyrAspLeuHisAspValLys 394
362 TCATGCGCCGAAATTTGACAGAGTACATCAGCAGAGGCTATGATCTCCATGATGTAAAA 421
395 GluTyrGlyGlnMetLeuLysAspAlaGlyPheValAspValLeuAlaGluAspArgThr 414
422 GCTTATGCGCAGATCTTGAAGGATGCTGTGTTTGTATCAGGTCAATTCAGAGGATCGAAC 481
415 GluGlnPheLeuArgValLeuArgLysGluLeuGluThrValGluLysGluLysAspVal 434
482 GATCAGTTTCATACAAAGTTCTGCAGCGTGAATTAGACAAAGTTGAGAGGACGAGTACC 541
435 PheLeuSerAspPheSerGluGluAspTyrAsnAspIleValGlyGlyTyr 451
542 TTCATCAGCGACTTTTCTCAGAGAGACTACAGAGATATCGTTGGGGGATGG 592

```

## RESULT 11

```

US-10-424-599-114653
: Sequence 114653, Application US/10424599
: Publication No. US20040031072A1
: GENERAL INFORMATION:
: APPLICANT: La Rosa Thomas J
: APPLICANT: Kovalic David K
: APPLICANT: Zhou Yihua
: APPLICANT: Cao Yongwei
: TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
: FILE REFERENCE: 38-21(53223)B
: CURRENT APPLICATION NUMBER: US/10/424,599
: CURRENT FILING DATE: 2003-04-28
: NUMBER OF SEQ ID NOS: 285684
: SEQ ID NO 114653
: LENGTH: 906
: TYPE: DNA
: ORGANISM: Glycine max
: FEATURE:
: OTHER INFORMATION: Clone ID: PAT_MBT3847_74541C.1
US-10-424-599-114653

```

## Alignment Scores:

```

Pred. No.: 5,02e-95 Length: 906
Score: 839.00 Matches: 158
Percent Similarity: 79.32% Conservative: 30
Best Local Similarity: 66.67% Mismatches: 49
Query Match: 33.36% Indels: 0
DB: 13 Gaps: 0

```

US-10-031-331B-40 (1-473) x US-10-424-599-114653 (1-906)

237 GluArgValPheGlyProGlyTyrValSerThrGlyGlyTyrGluThrThrLeuPhe 255  
1 GAGAGTGTGTTTGGTCAAGGTTTGTGAGCAGAGGAGGACTTGAACAACCAACCAATTC 60  
257 ValSerMetLeuAspPheGlyProGlyGlnLeuValLeuAspValGlyCysGlyLeuGly 276  
61 GTGGCAAGTTGGGCAATTAACCTGGCCGAGAGTACTGATGTTGGTGTGGTACTGGG 120  
277 GlyGlyAspPheTyrMetAlaGluThrPheAspValGluValValGlyPheAspLeuSer 296  
121 CGAGGTGCATTTATCATGGCAGAGAAATTTTGTATTTGAGGTTGTGGCATTTGACCTTCC 180  
297 ValAsnMetIleSerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaValGluPhe 316  
181 ATAAACAATATTTCTCTGGCAATTTGAACGTTGCTATTGGACTCAATGCTGTGTGAATTT 240  
317 GluValAlaAspCysThrThrIleAsnTyrProAspAsnSerPheAspValIleTyrSer 336  
241 GAGTGTGGGATTTGACATAAAACCAATTTCTGTGATACATTTGATGATATCTATTTCC 300  
337 ArgAspThrIleLeuHisIleGlnAspLysProAlaLeuPheArgSerPheTyrIleTyr 356  
301 CGTGACACATTTGATACATCAAAAGATAAGCCATCACTATTTCAGATCATTTTACAAATGG 360  
357 LeuLysProGlyGlyLysValLeuIleSerAspTyrCysAlaLysAlaGlyProProSer 376  
361 TTGAGCGTGGAGTACACTTCTGATTACTGATTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
377 ProGluPheAlaAlaTyrIleGlnArgGlyTyrAspLeuHisAspValLysGlyTyr 396  
421 TTAGGATATGCTGAGTATATTAAGAAAGGGGGGATATATATTTATCATGATGAAACATAT 480  
397 GlyGlnMetLeuLysAspAlaGlyPheValAspValLeuAlaGluAspArgThrGluGln 416  
481 TGTGGGATGCTGAGATGCTGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 540  
417 PheIleArgValLeuArgLysGluLeuGluThrValGluLysGluLysAspValPheIle 436  
541 TTATGAAACACTACACAGAGGTTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
437 SerAspPheSerGluGluAspTyrAsnAspIleValGlyGlyTyrAsnAspLysLeuArg 456  
601 GATGACTTCTCGAGAGAGACTACAAATGAAATTTACTGAAAGATGGAAGCCAGCAGATG 660  
457 ArgThrAlaLysGlyGluGlnArgTyrGlyLeuPheValAlaLysLys 473  
661 CGGGGGCGAGTGGTGAACAAATATGAGCTTGTTCATTCGCAAGAAAGAA 711

RESULT 12  
US-10-021-323-8285  
; Sequence 8285, Application US/10021323  
; Publication No. US20040123340A1  
; GENERAL INFORMATION:  
; APPLICANT: Deikman, Jill  
; APPLICANT: Peng, Paul C.C.  
; APPLICANT: Fincher, Karen L.  
; APPLICANT: Ziegler, Todd E.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(52274)B  
; CURRENT APPLICATION NUMBER: US/10/021,323  
; CURRENT FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: US 60/255, 619  
; PRIOR FILING DATE: 2000-12-14  
; NUMBER OF SEQ ID NOS: 17880  
; SEQ ID NO 8285  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Gossypium hirsutum  
; FEATURE: OTHER INFORMATION: Clone ID: LIB3828-008-Q1-K6-86  
US-10-021-323-8285

Alignment Scores:  
Pred. No.: 4,62e-95 Length: 601  
Score: 837.00 Matches: 162  
Percent Similarity: 88.38% Conservative: 13  
Best Local Similarity: 81.82% Mismatches: 23  
Query Match: 33.28% Indels: 1  
DB: 17 Gaps: 0  
US-10-031-331B-40 (1-473) x US-10-021-323-8285 (1-601)  
QY 254 LysGluPheValSerMetLeuAspLeuLysProGlyGlnLysValLeuAspValGlyCys 273  
DB 8 AAAAGATTGTGGGGAAGTTAGATCTTAAGCCCTGGCCAAAGTC-CTAGATGTTGGCTGT 66  
QY 274 GlyIleGlyGlyAspPheTyrMetAlaGluThrPheAspValGluValValGlyPhe 293  
DB 67 GCGATTGGTGGAGGTGACATTTATATGGCTGAGCAATTTGATGTTTCATGTTTGGGCAAC 126  
QY 294 AspLeuSerValAsnMetIleSerPheAlaLeuGluArgSerIleGlyLeuLysCysAla 313  
DB 127 GACCTCTCTGTAAACATGATATCTTGTCTTTGAACGAGCTACTGGACTGGAATGCTCA 186  
QY 314 ValGluPheGluValAlaAspCysThrLysIleAsnTyrProAspAsnSerPheAspVal 333  
DB 187 GTGGAAATTTGAAGTTGCTGATTGCCAAGAGGTTTATCCGACCAACAGTTTGTGTT 246  
QY 334 IleTyrSerArgAspThrIleLeuHisIleGlnAspLysProAlaLeuPheArgSerPhe 353  
DB 247 ATCTACACGCGTGACACTATTCTACATTCATGACAAACCTGCACCTGTTAGATCTTTC 306  
QY 354 TyrLysThrLeuLysProGlyGlyLysValLeuLysAspTyrCysAlaLysAlaGly 373  
DB 307 TACAATGTTTGAAGCCAGGAGGCAACTCTCTATAGTGATTTACTGCAAAAGTTCCAG 366  
QY 374 ProProSerProGluPheAlaAlaTyrIleGlnArgGlyTyrAspLeuHisAspVal 393  
DB 367 ACTCCATCCAGGAGTTTGTGCTGATATATCAACGACAGAGGCTATGATCTTCATGATGA 426  
QY 394 LysGluTyrGlyGlnMetLeuLysAspAlaGlyPheValAspValLeuAlaGluAspArg 413  
DB 427 AAATCATATGACAGATGCTTGAGAGTCTGTTTGTATGATGATTTCTTCACAGGATCGA 486  
QY 414 ThrGluGlnPheIleArgValLeuArgLysGluLeuGluThrValGluLysGlyLysAsp 433  
DB 487 ACCGATCAGTTCTTACAGTTCTCGGCGGTGAATTGAACCAAGTGGAGAAAGAGAGAT 546  
QY 434 ValPheIleSerAspPheSerGluGluAspTyrAsnAspIleValGlyGlyTyr 451  
DB 547 GCATTCATCTCTGACTTCTCTAAGGAGAGACTATGATGAAATAGTTGGTGGATGG 600

RESULT 13  
US-10-021-323-13961  
; Sequence 13961, Application US/10021323  
; Publication No. US20040123340A1  
; GENERAL INFORMATION:  
; APPLICANT: Deikman, Jill  
; APPLICANT: Peng, Paul C.C.  
; APPLICANT: Fincher, Karen L.  
; APPLICANT: Ziegler, Todd E.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(52274)B  
; CURRENT APPLICATION NUMBER: US/10/021,323  
; CURRENT FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: US 60/255, 619  
; PRIOR FILING DATE: 2000-12-14  
; NUMBER OF SEQ ID NOS: 17880  
; SEQ ID NO 13961  
; LENGTH: 586  
; TYPE: DNA  
; ORGANISM: Gossypium hirsutum  
; FEATURE: OTHER INFORMATION: Clone ID: LIB3828-008-Q1-K6-86  
US-10-021-323-8285

LOCATION: (1) (586)  
OTHER INFORMATION: unsure at all n locations  
OTHER INFORMATION: Clone ID: LIB3829-022-Q6-K6-H6  
US-10-021-323-13961

Alignment Scores:  
Pred. No.: 3 11e-90 Length: 586  
Score: 798.50 Matches: 150  
Percent Similarity: 85.56% Conservative: 10  
Best Local Similarity: 80.21% Mismatches: 26  
Query Match: 31.75% Indels: 1  
DB: 17 Gaps: 1

US-10-031-331B-40 (1-473) x US-10-021-323-13961 (1-586)

QY 106 PheSerAsnTTPLeuMetTyrLeuSerAspGluValGluAsnLeuValGluArg 125  
DB 26 TTCTCACTGGCTGTTGATGATATCTCTCANTAAAGAGTTGAGAA---TTTGGCCGAAGG 82  
QY 126 MetLeuLysTTPLeuLysProGlyGlyTyrIlePhePheArgGluSerCysPheHisGln 145  
DB 83 ATGCTCAAGTGGTTGAAGGTTGGTGACATATTTCTTCAGGAGTCTTGTCTTCATCA 142  
QY 146 SerGlyAspHisLysArgLysSerAsnProThrHisTyrArgGluProArgPheTyrThr 165  
DB 143 TCTGTGATGTCAGAGAAAGACCAACCACTCCTATCGAGGCCAGATTTTCACT 202  
QY 166 LysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSerTyrGluLeuSerLeu 185  
DB 203 AGGTCTTTAAAGATGCCAAGCACTGATGTTCTGGAATTCATTGAACTTTCTCTT 262  
QY 186 LeuSerCysLysCysIleGlyAlaTyrValArgAsnLysLysAsnGlnIleSer 205  
DB 263 GTTGGCTGCAAGTCATGAGCTTATGTTAAGAACAAAGAAATCAGAAATTCAGATTGC 322  
QY 206 TTPLeuTTPGlnLysValAspSerLysAspAspGlyPheGlnArgPheLeuAspThr 225  
DB 323 TGGATCTTCCAAAAGTTGTTTCAGATATGATAGGATTCAGGCTCTCTTGATTTCT 382  
QY 226 SerGlnTyrLysCysAsnSerIleLeuArgTyrGluArgValPheGlyProGlyTyrVal 245  
DB 383 GTTCAAGTCAAACTTAATAGCAATCTCCGTTATAGCGGTCTTGTGGACAGATATGTG 442  
QY 246 SerThrGlyGlyTyrGluThrLysGluPheValSerMetLeuAspLeuLysProGly 265  
DB 443 AGCACAGGAGGAATGAAACAAACAAAGAAATTTGTGGGAGTTAGATCTTAAGCTGGC 502  
QY 266 GlnLysValLeuAspValGlyCysGlyIleGlyGlyAspPheTyrMetAlaGluThr 285  
DB 503 CAAGAGTCTTATGATTTGGCTGTGGCATTTGGTGGAGTGACATTTATATGCTGAGGA 562  
QY 286 PheAspValGluValGly 292  
DB 563 TTTGATGTTTATGTTGGGC 583

## RESULT 14

US-10-021-323-17175/c  
Sequence 17175, Application US/10021323  
Publication No. US20040123340A1  
GENERAL INFORMATION:  
APPLICANT: Deikman, Jill  
APPLICANT: Feng, Paul C.C.  
APPLICANT: Fincher, Karen L.  
APPLICANT: Ziegler, Todd E.  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(52274)B  
CURRENT APPLICATION NUMBER: US/10/021,323  
CURRENT FILING DATE: 2001-12-12  
PRIOR APPLICATION NUMBER: US 60/255,619  
NUMBER OF SEQ ID NOS: 17880  
SEQ ID NO 17175

LENGTH: 694  
TYPE: DNA  
ORGANISM: Gossypium hirsutum  
FEATURE:  
OTHER INFORMATION: Clone ID: LIB3829-031-Q6-K6-F4-PPOLYA  
US-10-021-323-17175

Alignment Scores:  
Pred. No.: 2 11e-86 Length: 694  
Score: 769.00 Matches: 144  
Percent Similarity: 86.89% Conservative: 15  
Best Local Similarity: 78.89% Mismatches: 24  
Query Match: 20.58% Indels: 0  
DB: 17 Gaps: 0

US-10-031-331B-40 (1-473) x US-10-021-323-17175 (1-694)

QY 290 ValValGlyPheAspLeuSerValAsnMetIleSerPheAlaLeuGluArgSerIleGly 309  
DB 692 GTTGTGGCATCGATCTGTCCATACATATCAAAATCTTTGCTCTTGAAGAGCCATGCA 633  
QY 310 LeuLysCysAlaValGluPheGluValAlaAspCysThrLysIleAsnTyrProAspAsn 329  
DB 632 CTCGATTCTCTGTGGAATTTGAAGTTGCTGATTCGACTACAAAGACTTATCCGATAAC 573  
QY 330 SerPheAspValIleTyrSerArgAspThrIleLeuHisIleGlnAspLysProAlaLeu 349  
DB 572 AGCTTTGATGTTATCTACAGCCGTGACACGATCTACACATTCAGACAAAGCTTCGACTA 513  
QY 350 PheArgSerPheTyrLysTyrLeuLysProGlyGlyValLeuLysSerAspTyrCys 369  
DB 512 TTTAGATATTTCTACAAATGTTGAAGCCAGGGGCAAGTTCTTATAGTGTATTACTGC 453  
QY 370 LysLysAlaGlyProProSerProGluPheAlaIleTyrIleLysGlnArgGlyTyrAsp 389  
DB 452 AAAGCCGAGGAGCTTTCATGCCCGAATTTGACAGATACATCAAGCAGAGGCTATGAT 393  
QY 390 LeuHisAspValLysGluTyrGlyGlnMetLeuLysAspAlaGlyPheValAspValLeu 409  
DB 392 CTCATGATGTTAAAGCTTATGGCCAGATGCTTAAGATGCTGTTTGTATCAGTCAAT 333  
QY 410 AlaGluAspAspThrGluGlnPheIleArgValLeuArgLysGluLeuGluThrValGlu 429  
DB 332 GCAGAGATTCGACGATCATGATTCATACAAAGTTCTCAGCGGGAATTAGACAAAGTTGAG 273  
QY 430 LysGluLysAspValPheIleSerAspPheSerGluAspTyrAsnAspIleValGly 449  
DB 272 AAGGACAGGATACCTTCATCCGACTTTTCTCAGAGAGACTACGACGATATCGTTGG 213  
QY 450 GlyTyrAsnAspLysLeuArgAspThrAlaLysGlyGluGlnArgTyrGlyLeuPheVal 469  
DB 212 CGATCGAAGCCAAAGCTTAATAGACCAAGTTTGGGAGCAGAGATGGGCTCTTTTCAT 153  
QY 470 AlaLysLys 472  
DB 152 GCCAACAA 144

## RESULT 15

US-10-021-323-14357  
Sequence 14357, Application US/10021323  
Publication No. US20040123340A1  
GENERAL INFORMATION:  
APPLICANT: Deikman, Jill  
APPLICANT: Feng, Paul C.C.  
APPLICANT: Fincher, Karen L.  
APPLICANT: Ziegler, Todd E.  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(52274)B  
CURRENT APPLICATION NUMBER: US/10/021,323  
CURRENT FILING DATE: 2001-12-12  
PRIOR APPLICATION NUMBER: US 60/255, 619  
PRIOR FILING DATE: 2000-12-14

; NUMBER OF SEQ ID NOS: 17880  
; SEQ ID NO 14357  
; LENGTH: 525  
; TYPE: DNA  
; ORGANISM: Cossyplium hirsutum  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3829-033-Q1-K6-A5  
US-10-021-323-14357

Alignment Scores:  
Pred. No.: 1.79e-86 Length: 525  
Score: 768.00 Matches: 146  
Percent Similarity: 89.66% Conservative: 10  
Best Local Similarity: 83.91% Mismatches: 18  
Query Match: 30.54% Indels: 0  
DB: 17 Gaps: 0

US-10-031-331B-40 (1-473) x US-10-021-323-14357 (1-525)

```
QY 251 GluThrThrLysGluPheValSerMetLeuAspLeuLysProGlyGlnLysValLeuAsp 270
DB 4 GAACACACAAAGAAATTGTGGGAGGTAGATCTTARGCCTGGCCAAAGGTCCTAGAT 63
QY 271 ValGlyCysGlyIleGlyGlyLysPheTyrMetAlaGluThrPheAspValGluVal 290
DB 64 GTTGCTGTGGCATTTGTGGAGGTGACATTTATATGCTGAGGAATTTGATGTTTCATGTT 123
QY 291 ValGlyPheAspLeuSerValAsnMetIleSerPheAlaLeuGluArgSerIleGlyLeu 310
DB 124 GTGGGCAACGACCTCTCTGTAAACATATATCTTTGCTCTTGAACGAGCTACTGGACTG 183
QY 311 LysCysAlaValGluPheGluValAlaAspCysThrIleAsnTyrProAspAsnSer 330
DB 184 AAATGCTCAGTGGAAATTTGAAGTTGCTGATTCACCAAGAAAGCTTTATCCGGACACAGT 243
QY 331 PheAspValIleTyrSerArgAspThrIleLeuHisIleGlnAspLysProAlaLeuPhe 350
DB 244 TTGATGTTTACTACGCGGTGACACTATTCTACACATTCATCACAACCTGCACCTGTTT 303
QY 351 ArgSerPheTyrLysThrLeuLysPheProGlyGlyLysValLeuLysSerAspTyrCysLys 370
DB 304 AGATCTTTCTACAAATGTTGAAGCCAGGAGGCAACTCTCTATAAGTATGATTACTGCATA 363
QY 371 LysAlaGlyProProSerProGluPheAlaAlaTyrIleLysGlnArgGlyTyrAspLeu 390
DB 364 AGTTCCACAGACTCCATCCAGGAGTTTGTCTGAGTATATCAAGCAGAGAGGCTATGATCTT 423
QY 391 HisAspValLysGluTyrGlyGlnMetLeuLysAspAlaGlyPheValAspValLeuAla 410
DB 424 CATGATGTAATAATCATATGACAGATCTTGAGGATGCTGAGTATGATGATGATTTGCA 483
QY 411 GluAspArgThrGluGlnPheIleArgValLeuArgLysGlu 424
DB 484 GAGGATCGAACCGATCAGTTCTTACAGTTCTGCGGGGTGAA 525
```

Search completed: August 2, 2004, 05:04:18  
Job time : 651 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 2, 2004, 00:38:15 ; Search time 102 Seconds

Title: US-10-031-331B-40  
Perfect score: 2515  
Sequence: 1 HTVDITTEAMMLDSQSLD.....KLRTAKGRMGLFVAKKK 473  
2573.449 Million cell updates/sec

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_p2n.model -DEV=xih  
-Q=/cgn2\_1/USPTO\_spool/US10031331/runat\_26072004\_120352\_15000/app\_query.fasta\_1.647  
-DB=Issued Patents NA -QWTF=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=ext -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10031331 cgn2\_1 1 56 @runat\_26072004\_120352\_15000 -NCPU=6 -ICPU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DRV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:\*

1: /cgn2\_6/ptodata/2/ina/5A COMB seq:.\*  
2: /cgn2\_6/ptodata/2/ina/5B COMB seq:.\*  
3: /cgn2\_6/ptodata/2/ina/6A COMB seq:.\*  
4: /cgn2\_6/ptodata/2/ina/6B COMB seq:.\*  
5: /cgn2\_6/ptodata/2/ina/PTUS COMB seq:.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	405	16.1	286	4	US-09-313-294A-7492
2	363	14.4	280	4	US-09-313-294A-6753
3	347	13.8	274	4	US-09-313-294A-3443
4	330	13.1	260	4	US-09-313-294A-876
5	262	10.4	174	4	US-09-313-294A-5689
6	257.5	10.2	174	4	US-09-313-294A-1188
7	218.5	8.7	852	4	US-09-313-294A-1188
8	218.5	8.7	53500	4	US-09-313-294A-1188
9	210.5	8.4	954	4	US-09-313-294A-1188
10	201.5	8.0	6085	3	US-09-029-603-4
11	194.5	7.7	80161	3	US-09-036-987A-1
12	194.5	7.7	80161	3	US-09-370-700-1

C	13	194.5	7.7	80161	4	US-09-603-207-1	Sequence 1, Appli
	14	185.5	7.4	2122	3	US-09-029-603-1	Sequence 1, Appli
	15	184.5	7.3	750	4	US-09-134-000C-3308	Sequence 3308, Ap
	16	184.5	7.3	957	4	US-09-382-906A-1	Sequence 1, Appli
	17	181	7.2	828	4	US-09-266-965-35	Sequence 35, Appli
	18	175	7.0	1790	4	US-09-118-637A-3	Sequence 3, Appli
	19	174.5	6.9	1080	4	US-09-328-352-3879	Sequence 3879, Ap
	20	172	6.8	80161	3	US-09-036-987A-1	Sequence 1, Appli
	21	172	6.8	80161	3	US-09-370-700-1	Sequence 1, Appli
	22	172	6.8	80161	3	US-09-603-207-1	Sequence 1, Appli
C	23	167.5	6.7	997	4	US-08-956-171B-981	Sequence 981, App
	24	167.5	6.7	1369	4	US-09-439-554-29	Sequence 29, Appli
	25	165	6.6	771	4	US-09-328-352-3576	Sequence 3576, Ap
	26	163	6.5	1149	2	US-08-844-305-1	Sequence 1, Appli
	27	162	6.4	11466	4	US-08-956-171B-444	Sequence 444, App
C	28	158	6.3	984	4	US-09-252-991A-12173	Sequence 12173, A
	29	158	6.3	1254	4	US-09-252-991A-12420	Sequence 12420, A
C	30	156.5	6.2	777	4	US-09-134-001C-407	Sequence 407, App
	31	156	6.2	1400	3	US-09-041-718-1	Sequence 1, Appli
	32	155.5	6.2	1314	4	US-09-644-907B-1	Sequence 1, Appli
	33	154.5	6.1	888	4	US-09-540-236-1514	Sequence 1514, Ap
	34	154.5	6.1	954	4	US-08-489-039A-873	Sequence 873, App
	35	154.5	6.1	62809	4	US-09-536-002-32	Sequence 32, Appli
	36	153.5	6.1	22108	3	US-09-053-197A-3	Sequence 3, Appli
	37	153.5	6.1	22108	4	US-09-085-761A-3	Sequence 3, Appli
C	38	152	6.0	4403765	3	US-09-103-840A-2	Sequence 1, Appli
C	39	152	6.0	4411529	3	US-09-103-840A-1	Sequence 2534, Ap
	40	151	6.0	1173	4	US-09-489-039A-2534	Sequence 3925, Ap
	41	150	6.0	171	4	US-09-313-294A-3925	Sequence 913, App
	42	149	5.9	762	4	US-09-489-039A-913	Sequence 12098, A
	43	149	5.9	783	4	US-09-252-991A-2322	Sequence 2322, Ap
C	44	147.5	5.9	1131	4	US-09-252-991A-2322	Sequence 2513, Ap
	45	147.5	5.9	1215	4	US-09-252-991A-2513	

#### ALIGNMENTS

RESULT 1  
US-09-313-294A-7492  
; Sequence 7492, Application US/09313294A  
; Patent No. 6476212  
; GENERAL INFORMATION:  
; APPLICANT: Lalngudi, Raghunath V.  
; APPLICANT: Ito, Laura Y.  
; APPLICANT: Sherman, Bradley K.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR  
; FILE REFERENCE: PL-0017 US  
; CURRENT APPLICATION NUMBER: US/09/313,294A  
; CURRENT FILING DATE: 1999-05-14  
; NUMBER OF SEQ ID NOS: 7600  
; SOFTWARE: PERL Program  
; SEQ ID NO 7492  
; LENGTH: 286  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6476212 700381824H1  
US-09-313-294A-7492

Alignment Scores:  
Pred. No.: 1.65e-43  
Score: 405.00  
Length: 286  
Matches: 72  
Percent Similarity: 84.21%  
Conservative: 8  
Best Local Similarity: 75.79%  
Mismatch: 15  
Query Match: 16.10%  
Indels: 0  
Gaps: 0  
DB:

US-10-031-331B-40 (1-473) x US-09-313-294A-7492 (1-286)

QY 134 GlyTyrIlePheArgGluSerCysPheHISGLNserGlyAspHisLysArgLysSer 153  
DB 2 GGCCTATTTTCTTTAGAGNATCATGTTTCCACCAATCTGGAGATTCCTCAAAAGGAGTG 61



154 AsnProThrHisTyrArgGluProArgPheTyrThrLysAlaPheLysGluCysHisLeu 173  
Db 62 AACCCACACACTATCGAGACCAAGGTTTATACCAAGGATTTTAAAGAGCGCATTC 121  
Qy 174 GlnAspGlySerGlyAsnSerTyrGluLeuSerCysLysCysLeuGlyVala 193  
Db 122 TTTGATCAAGATGAGGTTTGGTTTGAACCTTCTTAGTGACCTGTAATGATTTGGGGCT 181  
Qy 194 TyrValArgAsnLysAsnGlnAenGlnLleSerTyrLeuTyrGlnLysValAspSer 213  
Db 182 TATGTCAAAACAGAGAATCAAAACCCAGATATGCTGTTATGGGAAAGGTAAATCA 241  
Qy 214 LysAspAspLysGlyPheGlnArgPheLeuAspThrSerGlnTyr 228  
Db 242 ACAGAGACAGAGATTTCAAGATTCCTGGACAGCTGCATAC 286  
RESULT 2  
US-09-313-294A-6753  
; Sequence 6753, Application US/09313294A  
; Patent No. 6476212  
; GENERAL INFORMATION:  
; APPLICANT: Lalgudi, Raghunath V.  
; APPLICANT: Ito, Laura Y.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR  
; FILE REFERENCE: PL-0017 US  
; CURRENT APPLICATION NUMBER: US/09/313,294A  
; CURRENT FILING DATE: 1999-05-14  
; NUMBER OF SEQ ID NOS: 7600  
; SOFTWARE: PERL Program  
; SEQ ID NO 6753  
; LENGTH: 280  
; TYPE: DNA  
; ORGANISM: Zea mays  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No. 6476212 700352341H1  
US-09-313-294A-6753  
Alignment Scores:  
Pred. No.: 4.71e-38 Length: 280  
Score: 363.00 Matches: 68  
Percent Similarity: 82.80% Conservative: 9  
Best Local Similarity: 73.12% Mismatches: 16  
Query Match: 14.43% Indels: 0  
Gaps: 4  
US-10-031-331B-40 (1-473) x US-09-313-294A-6753 (1-280)  
Qy 359 ProGlyGlyValLeuIleSerAspTyrCysLysAlaGlyProProSerProGlu 378  
Db 2 CCTGGGGCGCAGGTCCTTATCAGTCACTACTGCGAGGAGTCCTGGGAAACCATCAGAGAA 61  
Qy 379 PheAlaAlaTyrLysGlnArgGlyTyrAspLeuHisAspValLysGluTyrGlyGln 398  
Db 62 TTTCAGCGTACATTAAGCGAGAGGTTATGACCTACATCTGTGGAGGCTTATGCACAG 121  
Qy 399 MetLeuLysAspAlaGlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIle 418  
Db 122 ATGTGAAGAGTGCTGGTTTTCGTATGTCATCTGAGGATCGAATGATCATGTCCTT 181  
Qy 419 ArgValLeuArgLysGluLeuGluThrValGlnLysGluLysAspValPheIleSerAsp 438  
Db 182 GGTGTTTTAGATTAAGGATGAGTCTGAATTTGAAAGAACAGGACGATTTCTGTCTGAC 241  
Qy 439 PheSerGluGluAspTyrAsnAspIleValGlyTyr 451  
Db 242 TTCACCGAGGAGGATCAGTATGATATCGAAGCTTGG 280  
RESULT 3  
US-09-313-294A-3443  
; Sequence 3443, Application US/09313294A

Patent No. 6476212  
; GENERAL INFORMATION:  
; APPLICANT: Lalgudi, Raghunath V.  
; APPLICANT: Ito, Laura Y.  
; APPLICANT: Sherman, Bradley K.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR  
; FILE REFERENCE: PL-0017 US  
; CURRENT APPLICATION NUMBER: US/09/313,294A  
; CURRENT FILING DATE: 1999-05-14  
; NUMBER OF SEQ ID NOS: 7600  
; SOFTWARE: PERL Program  
; SEQ ID NO 3443  
; LENGTH: 274  
; TYPE: DNA  
; ORGANISM: Zea mays  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No. 6476212 700611594H1  
US-09-313-294A-3443  
Alignment Scores:  
Pred. No.: 5.52e-36 Length: 274  
Score: 347.00 Matches: 68  
Percent Similarity: 85.06% Conservative: 6  
Best Local Similarity: 78.16% Mismatches: 12  
Query Match: 13.80% Indels: 1  
Gaps: 0  
US-10-031-331B-40 (1-473) x US-09-313-294A-3443 (1-274)  
Qy 219 PheGlnArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArg 238  
Db 5 TTTTCAAGATTCCTGGACACGTCGCAATACAAACAGTGGGATATATCGTTACGAGCGT 64  
Qy 239 ValPheGlyProGlyTyrValSerThrGlyGlyTyrGluThrThrLysGluPheValSer 258  
Db 65 GTCTTTGCGAAGGTTTGTGCGACTGTGGAATCGAGACAAACAAAGGAATTTGTGGGC 124  
Qy 259 MetLeuAspLeuLysProGlyGlnLysValLeuAspValGlyCysGlyIleGlyGly 278  
Db 125 ATGCTCGATCTTAAACCGCGCGAGAGTACTTGTGATGTCGATGTCGATGTCGATGTCGCGC 184  
Qy 279 AspPheTyrMetAlaGluThrPheAspVal-GluValValGlyPheAspLeuSerValas 298  
Db 185 GACTTTTACATGCGCTGCNAACATATGATGTGCCATGTCTTGGTATTGATCTATCGGTGAA 244  
Qy 298 nMetIleSerPheAlaLeu 304  
Db 245 CATGTTTCATTTTGCATG 263  
RESULT 4  
US-09-313-294A-876  
; Sequence 876, Application US/09313294A  
; Patent No. 6476212  
; GENERAL INFORMATION:  
; APPLICANT: Lalgudi, Raghunath V.  
; APPLICANT: Ito, Laura Y.  
; APPLICANT: Sherman, Bradley K.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR  
; FILE REFERENCE: PL-0017 US  
; CURRENT APPLICATION NUMBER: US/09/313,294A  
; CURRENT FILING DATE: 1999-05-14  
; NUMBER OF SEQ ID NOS: 7600  
; SOFTWARE: PERL Program  
; SEQ ID NO 876  
; LENGTH: 260  
; TYPE: DNA  
; ORGANISM: Zea mays  
; NAME/KEY: misc\_feature

OTHER INFORMATION: Incyte ID No. 6476212 700549985H1  
NAME/KEY: unsure  
LOCATION: 24, 86  
OTHER INFORMATION: a, t, c, g, or other  
US-09-313-294A-876

Alignment Scores:  
Pred. No.: 8 29e-34 Length: 260  
Score: 330.00 Matches: 66  
Percent Similarity: 83.53% Conservative: 5  
Best Local Similarity: 77.65% Mismatches: 12  
Query Match: 13.12% Indels: 2  
DB: 4 Gaps: 0

US-10-031-331B-40 (1-473) x US-09-313-294A-876 (1-260)

QY 201 GlnGlnLileSerThrLeu-TpGlnLysValleAspSerLysAspLysGlyPheG1 220  
DB 3 CAAACCGAGATGCTGCTTANTGGGAAGGTAAATACACAGAGACAGAGATTTTCA 62  
QY 220 nArpPheLeuAsp-ThrSerGlnLysLysCysAsnSerLileLeuArgTyTGluArgValP 240  
DB 63 AAGATTCCTGCACACCTGCATNCCAAACCAAGTGGGATATTACGTTACGAGCGTGTCT 122  
QY 240 heGlyProGlyTyValSerThrGlyGlyTyThrGlyGluThrLysGluPheValSerMetL 260  
DB 123 TTGGTGAAGGTTTGTGAGCACTGTGGATCGACACACAAAGGATTTGTGGCGCATGC 182  
QY 260 euAspLeuLysProGlyGlnLysValleAspValGlyCysGlyLileGlyGlyAspP 280  
DB 183 TCGATCTTAACCGCGCCAGAAAGTACTTGATGTCCGATGTGGAAATTCGAGCGCGCACT 242  
QY 280 heTyMetAla 283  
DB 243 TTTACATGGCT 253

## RESULT 5

US-09-313-294A-5689  
Sequence 5689, Application US/09313294A  
Patent No. 6476212  
GENERAL INFORMATION:  
APPLICANT: Lalgudi, Raghunath V.  
APPLICANT: Ito, Laura Y.  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR  
FILE REFERENCE: PL-0017 US  
CURRENT APPLICATION NUMBER: US/09/313,294A  
CURRENT FILING DATE: 1999-05-14  
NUMBER OF SEQ ID NOS: 7600  
SOFTWARE: PERL Program  
SEQ ID NO 5689  
LENGTH: 174  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. 6476212 700350535H1  
US-09-313-294A-5689

Alignment Scores:  
Pred. No.: 3.03e-25 Length: 174  
Score: 262.00 Matches: 49  
Percent Similarity: 89.66% Conservative: 3  
Best Local Similarity: 84.48% Mismatches: 6  
Query Match: 10.42% Indels: 0  
DB: 4 Gaps: 0

US-10-031-331B-40 (1-473) x US-09-313-294A-5689 (1-174)

QY 234 LeuArgTyGluArgValPheGlyProGlyTyValSerThrGlyGlyTyThrGluThr 253  
DB 1 TTACGTTACGAGCGTGTCTTTGTGGAAGGTTTGTGAGCACTGTGTGGAATCGACACACA 60

QY 254 LysGluPheValSerMetLysAspLeuLysProGlyGlnLysValleAspValGlyCys 273  
DB 61 AAGGATTTTGGCGCATGCTCGATCTTAACCGCGCCAGAAAGTACTTGATGTGGATGT 120  
QY 274 GlyLleGlyGlyAspPheTyMetAlaGluThrPheAspValGluValVal 291  
DB 121 GGAATTGGAGCGCGCGACTTTTACATGGCTGCAAACTATGATGTCCATGTTCTT 174

## RESULT 6

US-09-313-294A-1188  
Sequence 1188, Application US/09313294A  
Patent No. 6476212  
GENERAL INFORMATION:  
APPLICANT: Lalgudi, Raghunath V.  
APPLICANT: Ito, Laura Y.  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR  
FILE REFERENCE: PL-0017 US  
CURRENT APPLICATION NUMBER: US/09/313,294A  
CURRENT FILING DATE: 1999-05-14  
NUMBER OF SEQ ID NOS: 7600  
SOFTWARE: PERL Program  
SEQ ID NO 1188  
LENGTH: 274  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. 6476212 700550521H1  
LOCATION: 256  
OTHER INFORMATION: a, t, c, g, or other  
US-09-313-294A-1188

## Alignment Scores:

Pred. No.: 2.53e-24 Length: 274  
Score: 257.50 Matches: 62  
Percent Similarity: 80.43% Conservative: 12  
Best Local Similarity: 67.39% Mismatches: 18  
Query Match: 10.24% Indels: 4  
DB: 4 Gaps: 0

US-10-031-331B-40 (1-473) x US-09-313-294A-1188 (1-274)

QY 48 ArgPheThrGlyGluLeuAlaGluLysAlaGlyGlnValleAlaLeuAspPheIleGlu 67  
DB 2 CGCTTTACTGGAGATCTGCGAAAGAGCTGGGCGACGTTCTGGCTCTAGACTTCATTGAA 61  
QY 68 SerAlaIleLysLysAsnGluValleAsnGlyHisTyLysAsnValLysPheMetCys 87  
DB 62 AGTGTGATTAAAGAAACCAAGCATATANTGGGCATCACAGACACATACCTTCAGGTGT 121  
QY 88 AlaAspValThrSerProThrLeuSerPheProHisSerLeuAspValIlePheSer 107  
DB 122 GCTGATGTGACATCTAACGACTTGAAGATTGAAT-AACTCTGTGATCTGATATTTTCA 180  
QY 108 AsnTrpLeuLeuMetTyTrpLeuSerAspGluGluValGluAsnLeuValGluArgMetLeu 127  
DB 181 AACTGGCTATTAAATGTACCT-TCAGATGAGGAGGTCCAAAGCTTTGGG-GAAATGCT- 237  
QY 128 LysTrpLeuLysProGlyGlyTyTrpIlePhePheArg 139  
DB 238 AATGGCTAANAAGTCGGGNGCCATATTTCTTTTGA 273

## RESULT 7

US-09-266-965-36  
Sequence 36, Application US/09266965  
Patent No. 6495348  
GENERAL INFORMATION:  
APPLICANT: Sherman, D  
APPLICANT: Mao, Y  
APPLICANT: Varoglu, M  
APPLICANT: He, M

; APPLICANT: Sheldon, P  
 ; TITLE OF INVENTION: Mitomycin biosynthetic gene cluster  
 ; FILE REFERENCE: 600.456US1  
 ; CURRENT APPLICATION NUMBER: US/09/266,965  
 ; EARLIER FILING DATE: 1999-03-12  
 ; EARLIER FILING DATE: 1996-08-19  
 ; EARLIER FILING DATE: 1994-10-06  
 ; EARLIER FILING DATE: 1993-10-07  
 ; EARLIER FILING DATE: 1993-10-07  
 ; NUMBER OF SEQ ID NOS: 145  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 36  
 ; TYPE: DNA  
 ; LENGTH: 852  
 ; ORGANISM: Streptomyces lavendulae  
 ; US-09-266-965-36

Alignment Scores:  
 Prod. No.: 2,12e-18 Length: 852  
 Score: 218.50 Matches: 68  
 Percent Similarity: 45.06% Conservative: 46  
 Best Local Similarity: 26.88% Mismatches: 92  
 Query Match: 8.69% Indels: 47  
 DB: 4 Gaps: 11

US-10-031-331B-40 (1-473) x US-09-266-965-36 (1-852)

```

Qy 236 TyrGluArgValPheGlyProGlyTyrValSerThrGly-----GlyTyr 250
Db 67 TAGGACCGGTTACCGCGCTGGGAGCGCTCTCCCTCGGAGAACCTGCACCTTCGGCTAC 126

Qy 251 -----GluThrThrlyseGluPheValSerMet 259
Db 127 TGGGACTCCCGCAGCAGCGTGGCGTGGCGGACCGCCAGCGGCTCACCGACATG 186

Qy 260 -----LeuAspLeuLysProGlyGlyValLeuAspValGlyCysGlyIle 275
Db 187 ATGGCGGAGCGGCTCGGCGTGGCGGCTCGGCGTGGCGGAGCGCGGCTCGCGGCTG 246

Qy 276 GlyGlyGlyAspPheTyrMetAlaGluThrPheAspValGluValGlyPheAspLeu 295
Db 247 GCGACCGCGGCTACGCGTACGCGTGGCGGCTGGCGGAGCGCATGTCCGCGGCTATC 306

Qy 296 SerValAsnMetIle-----SerPheAlaLeuGluArgSerIleGlyLeuLysCysAla 313
Db 307 AGCCATGACGAGTGTCTCGGCGCCACGCGCTGGCGGAGCGCGGCTCGCGGACCG 366

Qy 314 ValGluPheGluValAlaAspCysThrlyseIleAsnTyrProAspAsnSerPheAspVal 333
Db 367 GCGCGCTTCACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 426

Qy 334 IleTyrSerArgAspThrIleLeuHisIleGlnAspLysProAlaLeuPheArgSerPhe 353
Db 427 GTCATCGCCCTCGAATCGATCATCATCATCATCATCATCATCATCATCATCATCAT 486

Qy 354 TyrLysTrpLeuLysProGlyGlyValLeuLysSerAspTyrCysLysAlaGly 373
Db 487 GCGCGGCTGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 543

Qy 374 ProProSerProGluPheAlaAlaTyrIleLysGlnArgGlyTyrAspLeuHisAsp--- 392
Db 544 CCCTCGCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 594

Qy 393 -----VallyseGluTyrGlyGlnMetLeuLysAspAlaGly----- 404
Db 595 ATGATGACCGTGTGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 654

Qy 405 -----PheValAspValLeuAlaGluAspArgThrGluInPheIleArgValLeu 421
Db 655 CTGGAGGAGTTCCTCGACATC---AGCGACGAGACCTTGGAGAGACCTTTCAGGCTGCTC 711
  
```

```

Qy 422 ArgLysGluLeuGluThrValGluLysGluLysAspValPheIleSerAspPheSerGlu 441
Db 712 TCG-----GAGCGCATCAACTCTCGAAGCGAGAGCGCTGGAGACCGCAGCTTCGGCAG 762

Qy 442 Glu-----AspTyrAsnAspIleValGly 449
Db 763 GAGATGCTGAACCACTTCGACCCCGCGGACCTTCGTCGCGC 801
  
```

# RESULT 8

US-09-266-965-76/c  
 ; Sequence 76, Application US/09266965  
 ; Patent No. 6495348  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sherman, D  
 ; APPLICANT: Mao, Y  
 ; APPLICANT: Varoglu, M  
 ; APPLICANT: He, M  
 ; APPLICANT: Sheldon, P  
 ; TITLE OF INVENTION: Mitomycin biosynthetic gene cluster  
 ; FILE REFERENCE: 600.456US1  
 ; CURRENT APPLICATION NUMBER: US/09/266,965  
 ; EARLIER FILING DATE: 1999-03-12  
 ; EARLIER FILING DATE: 1996-08-19  
 ; EARLIER FILING DATE: 1994-10-06  
 ; EARLIER FILING DATE: 1994-10-06  
 ; EARLIER FILING DATE: 1993-10-07  
 ; NUMBER OF SEQ ID NOS: 145  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 76  
 ; TYPE: DNA  
 ; LENGTH: 53500  
 ; ORGANISM: Streptomyces lavendulae  
 ; US-09-266-965-76

Alignment Scores:  
 Prod. No.: 2,46e-15 Length: 53500  
 Score: 218.50 Matches: 68  
 Percent Similarity: 45.06% Conservative: 46  
 Best Local Similarity: 26.88% Mismatches: 92  
 Query Match: 8.69% Indels: 47  
 DB: 4 Gaps: 11

US-10-031-331B-40 (1-473) x US-09-266-965-76 (1-53500)

```

Qy 236 TyrGluArgValPheGlyProGlyTyrValSerThrGly-----GlyTyr 250
Db 7518 TAGGACCGGTTACCGCGCTGGGAGCGCTCTCCCTCGGAGAACCTGCACCTTCGGCTAC 7459

Qy 251 -----GluThrThrlyseGluPheValSerMet 259
Db 7458 TGGGACTCCCGCAGCAGCGTGGCGTGGCGGACCGCCAGCGGCTCACCGACATG 7399

Qy 260 -----LeuAspLeuLysProGlyGlyValLeuAspValGlyCysGlyIle 275
Db 7398 ATGGCGGAGCGGCTCGCATCGCGGCGGCTCGCGGCTCGCGGCTCGCGGCTCGCGG 7339

Qy 276 GlyGlyGlyAspPheTyrMetAlaGluThrPheAspValGluValGlyPheAspLeu 295
Db 7338 GCGACCGCGGCTAGGCGTACCGGCTCGCGGCTCGCGGCTCGCGGCTCGCGGCTCG 7279

Qy 296 SerValAsnMetIle-----SerPheAlaLeuLysSerIleGlyLeuLysCysAla 313
Db 7278 AGCCATGACGAGTGTCTCGGCGCCACGCGCTGGCGGAGCGCGGCTTCGCCGACCG 7219

Qy 314 ValGluPheGluValAlaAspCysThrlyseIleAsnTyrProAspAsnSerPheAspVal 333
Db 7218 GCGCGCTTCACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7159

Qy 334 IleTyrSerArgAspThrIleLeuHisIleGlnAspLysProAlaLeuPheArgSerPhe 353
Db 7158 GTCATCGCCCTCGAATCGATCATCATCATCATCATCATCATCATCATCATCATCAT 7099
  
```







Db 12899 AACTCGATGAGCTCGCGCGGTGAGAGCTGTGGGCTGTGGCAGCACTTACGACGACGCTAC 12840  
 QY 444 TyrAsnAspIle 447  
 Db 12839 TTGGGAGACATC 12828

RESULT 13  
 US-09-603-207-1/c  
 ; Sequence 1, Application US/09603207B  
 ; Patent No. 6521406  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baltz, Richard H  
 ; APPLICANT: Broughton, Mary C  
 ; APPLICANT: Crawford, Kathryn P  
 ; APPLICANT: Madduri, Krishnamurthy  
 ; APPLICANT: Madduri, Patti J  
 ; APPLICANT: Turner, Jan R  
 ; APPLICANT: Waldron, Clive  
 ; TITLE OF INVENTION: Bicosynthetic Genes For Spinosyn Insecticide  
 ; FILE REFERENCE: 50489 DIV1  
 ; CURRENT APPLICATION NUMBER: US/09/603,207B  
 ; CURRENT FILING DATE: 2000-06-23  
 ; EARLIER APPLICATION NUMBER: 09/370,700  
 ; EARLIER FILING DATE: 1998-03-09  
 ; NUMBER OF SEQ ID NOS: 39  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 1  
 ; LENGTH: 80161  
 ; TYPE: DNA  
 ; ORGANISM: Saccharopolyspora spinosa  
 US-09-603-207-1

Alignment Scores:  
 Pred. No.: 6.56e-12 Length: 80161  
 Score: 194.50 Matches: 74  
 Percent Similarity: 41.67% Conservative: 61  
 Best Local Similarity: 22.84% Mismatches: 137  
 Query Match: 7.73% Indels: 53  
 DB: 4 Gaps: 11

US-10-031-331B-40 (1-473) x US-09-603-207-1 (1-80161)

QY 172 HsLeuGlnAspGlySerGlyAsnSerTyrGluLeuSerLeuLeuSerCysLysCysIle 191  
 Db 13786 CACGAGGAGAGCCGACATCGCGGCTCGTACGACGCGGCTACCTGGAACGCAATTGGTC 13727  
 QY 192 GlyAlaTyrValArgAsn-----LysLysAsnGlnAsnGln----- 203  
 Db 13726 GGCCTCCACCTACCAACATCGCTTCTCGAGAGAGGCGTCAACCGCGAAGCGGC 13667  
 QY 204 ----IleSerTyrLeuTyrGlnLysValAsp-----Ser 213  
 Db 13666 GTTCTGCTTGGTGGTCCGAGGAGTCTGGACGACATATGCACTGGCCGACGCTGAACAGC 13607  
 QY 214 LysAspAspLysGlyPheGlnArgPheLeuAspThrSerGlnTyr-----Lys 229  
 Db 13606 GCGGAGGACGAGTGAACAGCAGAGGCGGAGCAATTCGACCGCGGATCAG 13548  
 QY 230 CysAsnSerIle-----LeuArgTyrGluArgValPheGlyProGlyTyrValSer 246  
 Db 13547 GTGAGTTCATTTGATGCGTGTGGCGACGGGCGTCCCTGACACCGGTTACTGGGGC 13488  
 QY 247 ThrGlyGlyTyr-----GluThr 252  
 Db 13487 ---GGCGGTATCGGGAGATGCGCGTGCCACACCGGTGCGGATGCTGCCGACCACTG 13431  
 QY 253 ThrLysGluPheValSerMetLeuAspLeuLysProGlyGlnLysValLeuAspValGly 272  
 Db 13430 ACCGACCTGTTCACGACAGCGCGCTCGTCCGCGGACGCGACCTGTTCGACCTGGGC 13371  
 QY 273 CysGlyIleGlyGlyAspPheTyrMetIleGluThrPheAspValGluValGly 292  
 Db 13370 TGGCGCAATGGGAGCCCCGTAGTCCGTGGCGGATGCGCCAGCGGCGTGTGAGTCAACCGGA 13311

QY 293 PheAspLeuSerValAsnMetIleSerPheAla-----LeuGluArgSerIleGlyLeu 310  
 Db 13310 ATCACCCTGAACCCGACGATCTCGCGCGCGCCACACGAGCTGCCAACGACGACGACTG 13251  
 QY 311 LysCysAlaValGluPheGluValAlaAspCysThrLysIleAsnTyrProAspAsnSer 330  
 Db 13250 GCGGCGAGTCTTGAAGTTCGATCTAGTCGACGCGCGCCAGCTCCCTACCCGACGGTTTC 13191  
 QY 331 PheAspValIleTyrSerArgAspThrIleLeuHisIleGlnAspLysProAlaLeuPhe 350  
 Db 13190 TTTCAGCCCGCATGGCGATCGCTCGTGTGCGAGATCTGGACACGCGCCCGCGATC 13131  
 QY 351 ArgSerPheTyrLysTrpLeuLysProGlyGlyCysValLeuLysSerAspTyrCysLys 370  
 Db 13130 CGCGAGGTCCACCGAATCTCGAACCAGCGCGCGCTTCTGCTCGGAGACATCATCACT 13071  
 QY 371 LysAlaGlyProProSerProGluPheAlaAlaTyrIleLysGlnArgGlyTyrAspLeu 390  
 Db 13070 CGGTTTCGACTCCCGGAA---GAGTACGCGCGGTT-----TGCACGCGCAGACGCC 13020  
 QY 391 HisAspValLysGluTyrGlyGlnMetLeuLysAspAlaGlyPheValAspValLeuAla 410  
 Db 13019 CATACCTTGAACAGCTTCACGCGCGCTGTGACGAGAGCGCGGTTCGAGATTCTCGAATC 12960  
 QY 411 GluAspArgThrGluGln-----PheIleArgValLeuArgLys 423  
 Db 12959 ACCGACCTCACGCGCACACAGCAGCGTGCATGCTCTCTGCTAGCTCGACGAGTTGCTCGG 12900  
 QY 424 GluLeuGluThrValGluLysGluLysAspValPheIleSerAspPheSerGluGluAsp 443  
 Db 12899 AAATCGATGAGCTCGCGCGGTGAGAGCTGTGGGCTGTGGCAGCTAC 12840

QY 444 TyrAsnAspIle 447  
 Db 12839 TTGGGAGACATC 12828

RESULT 14  
 US-09-029-603-1  
 ; Sequence 1, Application US/09029603  
 ; Patent No. 6210935  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Schupp, Thomas  
 ; APPLICANT: Engel, Natalie  
 ; APPLICANT: Bietenhader, Jurgen  
 ; APPLICANT: Toupet, Christine  
 ; APPLICANT: Pospiech, Andreas  
 ; TITLE OF INVENTION: Staurosporin Biosynthesis Gene Clusters  
 ; FILE REFERENCE: 4-20555/A/PCT  
 ; CURRENT APPLICATION NUMBER: US/09/029,603  
 ; CURRENT FILING DATE: 1998-03-20  
 ; EARLIER APPLICATION NUMBER: PCT/EP96/03643  
 ; EARLIER FILING DATE: 1996-08-19  
 ; NUMBER OF SEQ ID NOS: 11  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 1  
 ; LENGTH: 2122  
 ; TYPE: DNA  
 ; ORGANISM: Streptomyces longisporoflavus  
 ; FEATURE:  
 ; NAME/KEY: misc\_RNA  
 ; LOCATION: (1)..(2122)  
 ; OTHER INFORMATION: product = 2.1 kb region  
 US-09-029-603-1

Alignment Scores:  
 Pred. No.: 2e-13 Length: 2122  
 Score: 185.50 Matches: 68  
 Percent Similarity: 42.15% Conservative: 42  
 Best Local Similarity: 26.05% Mismatches: 92  
 Query Match: 7.38% Indels: 59  
 DB: 3 Gaps: 13



US-10-031-331B-40 (1-473) x US-09-029-603-1 (1-2122)

```
QY 236 TyrGluArgVal-----PheGlyPro----- 242
DB 895 TAGCAGCGCTGACGCTGAGCGCGATCAACGAGCGCTGTTCAACCCCAATGTGCACATC 954
QY 243 GlyTyrValSerThrGlyGlyTyrGluThrThrThrGlu----- 255
DB 955 GGTATTTGGACACCCCGGCTCGAGCGCCACCATCGAGGAGCGGATGACCGGCTCACC 1014
QY 256 -----PheValSerMetLeuAspLeuLeuPheProGlyGlnLysValLeuAspValGlyCys 273
DB 1015 GATGTGTTTCATCGAAGCGCTGACGCGTACCGCCACCTCCACCGCTCTCGACCTCGGCTGC 1074
QY 274 GlyLeuGlyGlyAspPheTyrMetAlaGluThrPheAspValGluValValGlyPhe 293
DB 1075 GGGTGGCGCGCGCGCTCAGGCTGTGGCGCGCACCGGGCAGCGGTTCACCGGATC 1134
QY 294 AspLeuSerValanMetIleSerPheAla-----LeuGluArgSerIleGlyLeuLys 311
DB 1135 AGCATCAGCAGGAGGAGATCAGGACCGCCCAACCGGCTGGCCGAGCGCGGCTCGCC 1194
QY 312 CysAlaValGluPheGluValAlaAspCysThrLysIleAsnTyrProAspAsnSerPhe 331
DB 1195 GACCGTGGCGGTTTCCAGCATGGCGAGCGCATGAACTGCCCTTCGCCGAGCGCTGTTTC 1254
QY 332 AspValIleTyrSerArgAspThrIleLeuHisIleGluAspLysProAlaLeuPheArg 351
DB 1255 GACCGCGTATGGCGCTGGATGATCTGCCATCCCGCAGCGGAGGAGGTTCACCC 1314
QY 352 SerPheTyrLysTrpLeuLysProGlyGlyValLeuIleSerAspTyrCysLysLys 371
DB 1315 GAGGTGTCGGGCTGCGCCCGGCGCGGATGCTCTCACCAGCATCTTCGAGCGC 1374
QY 372 AlaGlyProProSerProGluPheAlaAlaTyrIleLysGluArgGlyTyrAsp----- 389
DB 1375 -----CACCGCGCGAAGCG-----GTACGACCGCGGATCGACAGTTC 1416
QY 390 -----LeuHisAspValLysGluTyrGlyGlnMetLeuLysAsp 402
DB 1417 TGCGCGACCTGATGTGCACCGCGGACATCGACGACTAGCTGCGCTGTCGACCGC 1476
QY 403 AlaGly-----PheValAspValLeuAlaGluAspArgThrGluGlnPhe 417
DB 1477 TCGGCGCTGCGCTGCGGAGATGTCGAGTCT-----ACCGACGACGACC 1521
QY 418 IleArgValLeuArgLysGlu-----LeuGluThrValGluLysGluLysAspVal 434
DB 1522 ACGTGGCTTCGCGGAGAGATCGGAGGCTCGCGCGCTCGAG-----GAGCGCGCGGTG 1578
QY 435 PheIleSerAsp-----PheSerGluGluAspTyr-----AsnAspIleVal 448
DB 1579 OCCATGGACGAGGCGCAACTTCGCTTCGCGCAGCACTCTTCAGCGCGCTCGGCGTGGCG 1638
QY 449 Gly 449
DB 1639 GGC 1641
```

## RESULT 15

```
US-09-134-000C-3308
; Sequence 3308, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 3308
; LENGTH: 750
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-3308
```

```
Alignment Scores:
Pred. No.: 4,59e-14 Length: 750
Score: 184.50 Matches: 72
Percent Similarity: 43.06% Conservative: 49
Best Local Similarity: 25.62% Mismatches: 109
Query Match: 7.34% Indels: 51
DB: 4 Gaps: 12
```

US-10-031-331B-40 (1-473) x US-09-134-000C-3308 (1-750)

```
QY 209 GlnLysValAspSerLysAspAspLys-----GlyPheGlnArgPheLeu 223
DB 4 AGGAGAGTTTCAGATGAAGAAAGAAACAAATATGATGATAATATATTTTTCAAAAATACAGT 63
QY 224 AspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArgValPheClyProGly 243
DB 64 CAATGAGTCCGCTCCAG-----AAAGGA 87
QY 244 TyrValSerThrGlyGlyTyrGluThrThrLysGluPheValSerMetLeuAspLeuLys 263
DB 88 CTGGCTGGTGGCGGAGAAATGGGAGACTTTGAAAGAGATGCTACCT-----GATTTTAAG 141
QY 264 ProGlyGlnLysValLeuAspValGlyCysGlyIleGlyGlyAspPheTyrMetAla 283
DB 142 ---GCTAAGCGTGTCTTGATTTAGCATGCGCTATGATGCGCACTGATATATGCGCATG 198
QY 284 GluThrPheAspValGluValValGlyPheAspLeuSerValanMetIleSerPheAla 303
DB 199 GAAACCGTGTCTCTCTAGTAGGTGTGATTTCTCATATAATGCTCGAAGTAGCA 258
QY 304 LeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluAlaAspCysThrLys 323
DB 259 AAAGAAAGAACCCATTTTCCACAG-----ATTGAATATGATGCTGTGCCATAGAGAT 312
QY 324 IleAsnTyrProAspAsnSerPheAspValIleTyrSerArgAspThrIleLeuHisIle 343
DB 313 CTGGATTTCCAGAGGAGACTTTGATGATANTACTAGTTCCTTGGCTTCATTATGTA 372
QY 344 GlnAspLysProAlaLeuPheArgSerPheTyrLysTrpLeuLysProGlyGlyLysVal 363
DB 373 GCACGACTATGAGATTTTAATAAAAAAGATATATAGCATGCTGAAGGCTGGTGCATTTA 432
QY 364 LeuIleSerAspTyrCysLysLysAlaGlyProProSerProGluPheAlaAla 381
DB 433 GTTTTACA-----GTTGAACATCTCTGTTTACTGCTCATGGA 471
QY 382 -----TyrIleLysGlnArgGlyTyrAspLeuHisAspValLysGlu----- 395
DB 472 ACACAGACTGTTATATTAACGAAAGGAGAAATACTGATTCATTCGCCGTGGACAATPAT 531
QY 396 TyrGlyGlnMetLeuLysAspAlaGlyPheValAspValLeuAlaGluAspArgThrGlu 415
DB 532 TATTATGAGGCGAAACGACAGCATGATGTTTG-----GAAGAAAGGTGACA 579
QY 416 GlnPheIleArgValLeuArgLysGluLeuGluThrValGluLysGluLysAspVal--- 434
DB 580 AAATATCATAGACACTGACCACATATCTAATACACTGCTTTCMAATAGTTTTTATATA 639
QY 435 -----PheIleSerAspPheSerGluGluAspTyrAsnAspIleValGlyTyrAsn 452
DB 640 AATCAGATTGTGGAGCCACGCGCCAGAGACATGATGATATT---CCGGGATGCGG 696
QY 453 AspLysLeuArgArgThrAlaLysGlyGluGlnArgGlyPheValLysLysLys 472
DB 697 GATCAATGCGACGCGCCCATGATG-----CTGATTGTTATCGCAAAA 738
QY 473 Lys 473
```

Tue Aug 3 10:37:37 2004

us-10-031-331b-40.rni

Page 10

DB 739 AAG 741

Search completed: August 2, 2004, 04:40:06  
Job time : 219 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 1, 2004, 23:10:34 ; Search time 544 Seconds

(without alignments)  
3693.746 Million cell updates/sec

Title: US-10-031-331b-40

Perfect score: 2515

Sequence: 1 HTVDLTIEAMWLDQSASDLD.....KLRRATKGEQRMGLFVAKKK 473

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 04

Maximum Match 1004  
Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_p2n.model -DRV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US10031331/runat\_26072004\_120351\_14972/app\_query.fasta\_1.647  
-DB=N Geneseq 29Jan04 -QMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOPECL=0  
-LIST=45 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi  
-MODE=LOCAL -OUTFWT=pcp -NORM=ext -HEAPSIZ=500 -THR\_MIN=0 -ALIGN=15  
-USER=US10031331 -CGCN\_1\_470 -runat\_26072004\_120351\_14972 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DRV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 29Jan04:.\*  
1: geneseqn1980s:.\*  
2: geneseqn1990s:.\*  
3: geneseqn2000s:.\*  
4: geneseqn2001as:.\*  
5: geneseqn2001bs:.\*  
6: geneseqn2002as:.\*  
7: geneseqn2003as:.\*  
8: geneseqn2003bs:.\*  
9: geneseqn2003cs:.\*  
10: geneseqn2004s:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	2515	100.0	1602	AAF74206	DNA encod
2	2204	87.6	2235	AAD16797	Spinach P
3	1829.5	72.7	1962	AAH44262	Physcomit
4	1270	50.5	1120	AAD16798	Spinach P
5	905	36.0	636	ACL18332	DNA clone
6	828	32.9	605	ACL25710	DNA clone
7	804.5	32.0	667	ACL18333	DNA clone
8	785	31.2	611	ACL25711	DNA clone

C 9	769	30.6	777	8	ACL19405	DNA clone
C 10	767.5	30.5	575	8	ACL18347	DNA clone
C 11	751	29.9	566	8	ACL19390	DNA clone
C 12	747.5	29.7	650	8	ACL18342	DNA clone
C 13	732	29.1	622	8	ACL18325	DNA clone
C 14	719	28.6	487	9	ADE82005	Arabidops
C 15	697	27.7	589	8	ACL25708	DNA clone
C 16	682.5	27.1	517	8	ACL18349	DNA clone
C 17	661	26.3	736	8	ACL18395	DNA clone
C 18	658	26.2	619	4	AAH44246	Physcomit
C 19	657.5	26.1	587	8	ACL18344	DNA clone
C 20	647	25.7	671	8	ACL19401	DNA clone
C 21	646	25.7	738	8	ACL18354	DNA clone
C 22	641	25.5	712	8	ACL18373	DNA clone
C 23	639	25.4	727	8	ACL18362	DNA clone
C 24	633	25.2	676	8	ACL18366	DNA clone
C 25	632	25.1	648	8	ACL19403	DNA clone
C 26	629	25.0	663	8	ACL19412	DNA clone
C 27	623	24.8	621	8	ACL18331	DNA clone
C 28	617	24.5	659	8	ACL18389	DNA clone
C 29	612	24.3	624	8	ACL19409	DNA clone
C 30	612	24.3	649	8	ACL19408	DNA clone
C 31	612	24.3	652	8	ACL19407	DNA clone
C 32	611	24.3	623	8	ACL18326	DNA clone
C 33	610	24.3	696	8	ACL18382	DNA clone
C 34	604	24.0	697	8	ACL18372	DNA clone
C 35	603	24.0	681	8	ACL18374	DNA clone
C 36	601.5	23.9	527	8	ACL18334	DNA clone
C 37	598.5	23.8	488	8	ACL18346	DNA clone
C 38	594	23.6	436	8	ACL18323	DNA clone
C 39	592	23.5	692	8	ACL18391	DNA clone
C 40	590	23.5	684	8	ACL18377	DNA clone
C 41	589	23.4	477	8	ACL25703	DNA clone
C 42	585.5	23.3	536	8	ACL18350	DNA clone
C 43	578	23.0	695	8	ACL18380	DNA clone
C 44	574	22.8	631	8	ACL18356	DNA clone
C 45	574	22.8	673	8	ACL18378	DNA clone

ALIGNMENTS

RESULT 1

AAF74206  
ID AAF74206 standard; DNA; 1602 BP.  
AC AAF74206;  
XX  
DT 06-AUG-2003 (revised)  
DT 02-MAY-2001 (first entry)  
XX  
DE DNA encoding environmental stress tolerant protein SEQ ID 39.  
XX  
KW Environmental stress resistance; salt; heat; desert; transgenic plant;  
KW ds.  
XX  
OS Suaeda japonica.  
XX  
PN WO200106006-A1.  
XX  
PD 25-JAN-2001.  
XX  
PF 19-JUL-2000; 2000WO-JP004862.  
XX  
PR 19-JUL-1999; 99JP-00235910.  
PR 24-MAR-2000; 2000JP-00085377.  
XX  
(NISC-) JAPAN SCI & TECHNOLOGY CORP.  
XX  
PI Yamada A, Ozeki Y, Saito T.  
XX  
DR WPI; 2001-147355/15.  
DR P-PSDB; AAB80627.

PT Screening method to obtain DNA encoding environmental stress resistance  
 PT factor, useful for producing transgenic plants resistant to environmental  
 PT stress.

XX Claim 65: Page 123-125; 167pp; Japanese.

XX Polynucleotide sequences AAF74187 - AAF74218 encode proteins AAB80608 -  
 CC AAB80639, which impart environmental stress resistance. The invention  
 CC relates to a method for identifying DNA encoding proteins imparting  
 CC environmental stress resistance. The method comprises inserting cDNA from  
 CC a library originating in a salt-resistant organism into a host cell,  
 CC culturing the transformants under conditions in which the untransformed  
 CC host does not grow well, and selecting for viable clones. The method is  
 CC useful for obtaining DNA encoding environmental stress resistance  
 CC factors. The DNA encoding proteins conferring environmental stress  
 CC resistance, can be used in the production of plants resistant to  
 CC environmental stress, which can be cultivated in unfavourable  
 CC environments such as deserts, salt damaged ground, cold regions and the  
 CC oceans. They can be used for increasing the area of land covered by green  
 CC plants, and desert greening and afforestation, in order to counter the  
 CC effects of the increase in atmospheric carbon dioxide concentration. PCR  
 CC primers AAF74219 and AAF74220 are used in an example illustrating the  
 CC method of the invention. (Updated on 06-AUG-2003 to correct OS field.)

XX Sequence 1602 BP; 454 A; 279 C; 383 G; 486 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.: 1,3e-263 Length: 1602  
 Score: 2515.00 Matches: 473  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0

US-10-031-331b-40 (1-473) x AAF74206 (1-1602)

Qy 1 HistrValaspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp 20  
 Db 1 CACACCGTGTGATTAACCTTGAGCTATGATGCTGATCTTCAAGCTTCTGATCTTGAC 60  
 Qy 21 LysGluGluArgProGluIleLeuSerMetLeuProLeuGluGluLysCysLeuLeu 40  
 Db 61 AAGAAGAAAGCGCTGAGATTTCTTCAATGCTTCCGCTCTTGAAGGAAATGCTCTTG 120  
 Qy 41 GluLeuGluValacIlyleGlyArgPheThrGlyGluLeuAlaGluLysAlaGlyGlnVal 60  
 Db 121 GAACCTGGGCTGGTATTGGTGGTTTACGTGATTTGGCTGAGAAAGCTGGCCAGGTT 180  
 Qy 61 IleAlaLeuaspPheileGluSerAlaIleLysLysAsnGluValIleAsnGlyHisTyr 80  
 Db 181 ATTGCTCTGGATTTCAATTGAGAGTGCTATCAAGAAAGATGAAGTAATCAATGGGCACTAC 240  
 Qy 81 LysAsnValLysPheMetCysAlaaspValThrSerProThrLeuSerPheProHis 100  
 Db 241 AAAAATGTCAGTTTATGTGTGCTGATGCTACTTCTCCACTCTCAGTTTCCCAACCAT 300  
 Qy 101 SerLeuaspValIlePheSerAsnTrpLeuLeuMetTyrLeuSerAspGluGluValGlu 120  
 Db 301 TCATTGGATGATATTTCCATTTGGTTACTCATGTATCTTTCTGATGAGAGGTGGAA 360  
 Qy 121 AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyLysTrpIlePhePheArgGlu 140  
 Db 361 AATTGGTTGAAAGAAATGTTGAAATGGTTGAAGCCAGGGGGTTACATTTCTTCAGAGAA 420  
 Qy 141 SerCysPheHisGluSerGlyAspHisLysArgLysSerAsnProThrHisTyrArgGlu 160  
 Db 421 TCTGTGTTCCATCTCGGGATCACAACGCAAAAGCAATCCCAACCACTACCTCGAA 480  
 Qy 161 ProArgPheTyrThrIlyalalaphLysGluCysHisLeuGlnAspGlySerGlyAsnSer 180  
 Db 481 CCTAGGTTCTACACTAAGGCCCTTCAAGAGTGTCAATTGCAAGATGATGATCTCGAACTCT 540  
 Qy 181 TyrGluLeuSerLeuLeuSerCysLysCysValleGlyAlaTyrValArgAsnLysLysAsn 200

Db 541 TATGAGCTCTCCCTACTTACTGTCAAATGTATTGGAGCTTATGTGCAGAAACAGAAAC 600  
 Qy 201 GlnAsnGlnIleSerTrpLeuTrpGlnLysValaspSerLysAspAspLysGlyPheGln 220  
 Db 601 CAGAACCAAGATTAGTTGGTGGTGGCAAAAGTTGATCTTAAGGATGATAAGGGGTTCCAG 660  
 Qy 221 ArgPheLeuaspThrSerGlnTrpLysCysAsnSerIleLeuArgTyrGluArgValPhe 240  
 Db 661 CGATTTCTGGATACTCCAGTCAAGTGTATAGCATCTCCGATATGAGCGTGTATT 720  
 Qy 241 GlyProGlyTyrValSerThrGlyGlyTyrGluThrLysGluPheValSerMetLeu 260  
 Db 721 GGCCTCGTATTATGTAGCACTCGAGGATATGAACCAACCAAGAGCTTTGTGTCAATGCT 780  
 Qy 261 AspLeuLysProGlyGlnLysValLeuaspValGlyCysGlyIleGlyLysValaspPhe 280  
 Db 781 GACTTGAAGCTCGCCAGAGGCTCTGGATGTTGGTGGAAATTGGTGGAGGTGACTTT 840  
 Qy 281 TyrMetAlaGluThrPheaspValGluValaspValGlyPheaspLeuSerValAsnMetile 300  
 Db 841 TACATGGCGGAGACCTTGGATGTGGAGTGTGGATTTGGATTTCTCTCCTTATATGATT 900  
 Qy 301 SerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluValalasp 320  
 Db 901 TCTTTGCTTGGCTTGGAGCTTCTTATTGGGCTTAAATGTGTGTTGGTGGAGGTAGCAGT 960  
 Qy 321 CysThrLysIleAsnTyrProaspAsnSerPheaspValIleTyrSerArgaspThrile 340  
 Db 961 TGCAACAGATTAACCTACCTGATACCTTTTGTGATGTCATCTATAGCGCTGACACCAT 1020  
 Qy 341 LeuHisIleGlnAspLysProAlaLeuPheArgSerPheTyrLysTrpLeuLysProGly 360  
 Db 1021 CTGCATATTCAGGCAAGCCCTGCGTGTGTTAGATCTCTTCAAAATGGTTGAAGCAGGA 1080  
 Qy 361 GlyLysValLeuIleSerAspTyrCysLysLysAlaGlyProProSerProGluPheAla 380  
 Db 1081 GGTAAAGTTCTAAATCAGTGATTACTGCAGAAAGCTGGTCCACCTCCTCAATTCGCC 1140  
 Qy 381 AlaTyrIleLysGlnArgGlyCysAspLeuHisAspValLysGluTyrGlyGlnMetLeu 400  
 Db 1141 GCTTACATTAGCAGAGGGGATGATGATCTCATGTATGAAGGATATGGGCAAGTCTT 1200  
 Qy 401 LysAspAlaGlyPheValaspValLeuAlaGluaspArgThrGluIleArgVal 420  
 Db 1201 AAAGATGCTGGATTTGTGATGTTCTTCCGAGGATAGAACTGAGCAGTTCATTGAGTT 1260  
 Qy 421 LeuArgLysGluLeuGluThrValGluLysGluLysaspValPheIleSeraspPheSer 440  
 Db 1261 CTACGGAGGAACTAGAGACTTGTGAGAGAAAGGATGTGTTTCATTAGTTCATTCTCT 1320  
 Qy 441 GluGluaspTyrAsnaspIleValGlyIleTrpAsnaspLysLeuArgThrAlaLys 460  
 Db 1321 GAGGAGGATTAACATGACATTTGGAGGTTGGAAATGATTAAGTTGCGAGGACTGCCAG 1380  
 Qy 461 GlyGluGlnArgTrpGlyLeuPheValAlaLysLysLys 473  
 Db 1381 GGTGAGCAACGATGGGCTGTGTTGTTGCCAAGAAGAAG 1419

#### RESULT 2

AAD16797 standard; DNA; 2235 BP.

XX AAD16797;

XX 29-NOV-2001 (first entry)

XX Spinach PEAMT DNA.

XX Spinach: S-adenosyl-L-methionine; phosphoethanolamine N-methyltransferase;  
 KW PEAMT; cellular intermediate; phospho-dimethyl ethanolamine; choline;  
 KW phosphomono-methyl ethanolamine; phosphocholine; phosphatidylcholine;  
 KW glycine betaine; choline-O-sulphate; lipid content alteration;

osmotic stress tolerance; nutritional value; transgenic plant;  
cryoprotectant; ds.

Spinacia oleracea.

Key Location/Qualifiers

CDS 254..1738

/\*tag= a

/product= "Spinach PEAMT protein"

W0200168870-A2.

20-SRP-2001.

15-MAR-2001; 2001WO-US008352.

15-MAR-2000; 2000US-00525885.

(UYFL ) UNIV FLORIDA.

(UYCA-) UNIV CARNEGIE MELLON.

Hanson AD, Nuccio ML, Henry SA;

WPI; 2001-565796/63.

P-PSDB; AAE09760.

New plant S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase polypeptides; useful for modulating the levels of cellular intermediates such as phosphodimethylethanolamine and for altering the lipid content in plants cells.

Claim 18; Page 108-109; 158pp; English.

The present sequence is a DNA encoding spinach S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase (PEAMT). The PEAMT sequences are useful for modulating the levels of cellular intermediates such as phosphodimethylethanolamine, phosphomono-methylethanolamine, choline, betaine. They are useful for altering the lipid content in plant cells. The polynucleotides are also useful for improving the osmotic stress tolerance of a plant and increasing the cryoprotectant properties of a plant. The present invention also relates to methods and compositions comprising PEAMT used for generating transgenic plants with increased nutritional value

Sequence 2235 BP; 623 A; 399 C; 516 G; 697 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,61e-229	Length:	2235
Score:	2204.00	Matches:	408
Percent Similarity:	93.64%	Conservative:	34
Best Local Similarity:	86.44%	Mismatches:	30
Query Match:	87.63%	Indels:	0
DB:	5	Gaps:	0

US-10-031-331B-40 (1-473) x AAD16797 (1-2235)

Qy	1	HistHrValAspLeuThrileGluAlaMetMetLeuAspSerGlnAspSerLeuAsp	20
Db	317	CACCTCTGTTGATTTGACCTGTTGAGCTATGATGCTTCAATTCACAGCTTCAGATCTTGAC	376
Qy	21	LysGluGluArgProGluLeuLeuSerMetLeuProLeuGluGlyLysCysLeuLeu	40
Db	377	AAAGTCGAGCGACCTGAGTACTTCCATGCTCCACCTTATGAGGAAAGTCTGCTTA	436
Qy	41	GluLeuGluValGlyLeuGlyArgPheThrGlyGluLeuAlaGluValGlyGlnVal	60
Db	437	GAACCTCGGTGCTGTTGTCGCTTTTACTGTTGAAATGTCGCGAGAAAGCTAGCCAGGTC	496
Qy	61	IleAlaLeuAspPheileGluSerAlaIleLysLysAsnGluValIleAsnGlyHisTyr	80
Db	497	ATCGCTCTGCAATTCATTCAGAGTGTTATAAGGAAGAAATGAAAGCAATAATGGCATTAC	556

Qy	81	LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProHis	100
Db	557	AAATAATGTAAGTTTATGCTGCTGATGACATCTCCAAAGTCTCAACATTTCCACCAAT	616
Qy	101	SerLeuAspValIlePheSerAsnTrpLeuLeuMetTyrLeuSerAspGluGluValGlu	120
Db	617	TCCGTGGATATCATATTTCTCAATTTGGCTACTCATGTATCTTTCTGATGAAGGGTTGAG	676
Qy	121	AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyGlyTyrIlePhePheArgGlu	140
Db	677	CGTCTGGTTGAAGGATGTTGAAATGTTGAAGCCAGGAGGATACATTTCTTCAGAGAA	736
Qy	141	SerCysPheHisGlnSerGlyAspHisLysLeuSerAsnProThrHisTyrArgGlu	160
Db	737	TCTTGTGTTTTCATCATCAGGAGATCAACAGCGAAGCAATCCACCCACCTACCGTGA	796
Qy	161	ProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSer	180
Db	797	CCTAGGTTCTACACCAAGATCTTCAAGAAATGCATATGCAAGATGATTTCTGGGAATCC	856
Qy	181	TyrGluLeuSerLeuLeuSerCysLysCysIleGlyAlaTyrValArgAsnLysLysAsn	200
Db	857	TATGAGCTCTCCCTAATTTGGCTGCANATGTTATGGAGCTTATGTCAAGAGCAAGAGAT	916
Qy	201	GlnAsnGlnIleSerTrpLeuTrpGlnLysValAspSerLysAspAspLysGlyPheGln	220
Db	917	CAGAACCAAGATATAGCTGTTATGGCAGAAAGTTGATTCAGAGGATCAACAGGGTTCCAG	976
Qy	221	ArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArgValPhe	240
Db	977	CGATTTCTGGATTTCTAGTCAATCAAGTTTAAACAGCATACTGCTTATGAGGCTGTTT	1036
Qy	241	GlyProGlyTyrValSerThrGlyGlyTyrGluTrpThrLysGluPheValSerMetLeu	260
Db	1037	GCTCTGTTGATGTTAGTACCGAGAGCTCGAAACCAACCAAGAGGTTGTATCAAGCTT	1096
Qy	261	AspLeuLysProGlyGlnLysValLeuAspValGlyCysGlyIleGlyLysGlyAspPhe	280
Db	1097	GACTTGAAGCCTGGCCAGAAAGCTCTAGATGTGGTGTGGCATAGGTGAGGTGATTTT	1156
Qy	281	TyrMetalGluThrPheAspValGluValGlyPheAspLeuSerValAsnMetIle	300
Db	1157	TACATGGCAGAGAACTATGATGTTGAGTGTGGATTTGATCTCTCCATTATATGAT	1216
Qy	301	SerPheAlaLeuGluArgSerIleGlyLysCysAlaValGluPheGluValAlaAsp	320
Db	1217	TCTTTTGGCTTTGAGCGCTCAATTTGGCTCAATGCTGTTGAGTTGAGGTGGCAGAT	1276
Qy	321	CysThrLysIleAsnTyrProAspAsnSerPheAspValIleTyrSerArgAspThrIle	340
Db	1277	TGCACCAAGAAAGATTAACCTCGAAACTCTTTTATGATGTCATCTACAGCGGTGATACCAT	1336
Qy	341	LeuHisIleGlnAspLysProAlaLeuPheArgSerPheTyrLysTrpLeuLysProGly	360
Db	1337	CTGCATATTGAGGACAAACCTGCTTTATTAGATCTTCACCAATGTTGAAACCTGGA	1396
Qy	361	GlyLysValLeuLysSerAspTyrCysLysLysAlaGlyProProSerProGluPheAla	380
Db	1397	GGCAAGTCTTATTAGTACTACTGTAAAGAGTCTGTTGATACCTTCAGCTGCAATTTGCT	1456
Qy	381	AlaTyrIleLysGlnArgGlyTyrAspLeuHisAspValLysGluTyrGlyGlnMetLeu	400
Db	1457	GCATACATCAGGACAGAGGGATATGATCTCCAGATGTCAGAGCATATGGCAAGATGCTT	1516
Qy	401	LysAspAlaGlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIleArgVal	420
Db	1517	AAAGATGCTGGAATGCTTGGAGTTATGCTGAGATAGGACTGACCACTTCACTCAAGTT	1576
Qy	421	LeuArgLysGluLeuGluThrValGluLysGluLysAspValPheIleSerAspPheSer	440
Db	1577	CTGCAGAGGAAGTATAGTCTTGAACAGGAGAGAGATGACTTCACTTATGATGATTTCTCT	1636
Qy	441	GluGluAspTyrAsnAspIleValGlyGlyTyrPheAsnAspLysLeuArgThrAlaLys	460

Db 1637 GAGGAGGATTATACGACATAGTTGATGGTTGGAGGCCAGCTGGTGAGGACTACAGAG 1696  
 QY 461 GYGLUGLNGAGTTPGLYLeuPheValAlaLysLys 472  
 Db 1697 GGTGAGCAACATGGGGTTTGTTCATTTGCCAGAAA 1732

## RESULT 3

AAH44262

ID AAH44262 standard; DNA; 1962 BP.

XX

AC AAH44262;

XX

DT 21-SEP-2001 (first entry)

XX

DE Physcomitrella patens 78\_ppprot1\_092\_e12-260-rev gene.

XX

KW Tocopherol and carotenoid metabolism related protein; TCNRP; synthesis;  
 KW Physcomitrella patens; moss; algae; microorganism; fungus; plant;  
 KW identification; genome mapping; modulation; evolutionary study;  
 KW cellular production; fine chemical; db.

XX

OS Physcomitrella patens.

XX

XX WO200144276-A2.

XX

XX 21-JUN-2001.

XX

XX 14-DEC-2000; 2000WO-EP012698.

XX

XX 16-DEC-1999; 99US-0171121P.

XX

XX (BADI ) BASF PLANT SCI GMBH.

XX

PI Lerchl J, Renz A, Ehrhardt T, Reindl A, Cirpus P, Bischoff F;  
 PI Frank M, Freund A, Duwenig E, Schmidt R, Reski R, Badur R;

XX

XX WPI, 2001-398121/42.

XX

XX P-PSDB; AAB99889.

XX

PT Tocopherol and carotenoid metabolism related protein (TCNRP), used to  
 PT produce fine chemicals, is isolated from mosses, algae, microorganisms,  
 PT fungi, plants, or their fragments.

XX

XX Claim 6; Page 112-113; 123pp; English.

XX

CC The present invention describes isolated tocopherol and carotenoid  
 CC metabolism related proteins (TCNRP) (I) from mosses or algae,  
 CC microorganisms or fungi, plants, or its fragments. (I) can be used as  
 CC enzymes in the production of fine chemicals or in the metabolism of  
 CC tocopherols and carotenoids. (I) also assist in transmembrane transport.  
 CC The fine chemicals that can be produced include lipids, fatty acids,  
 CC vitamins, cofactors, enzymes, amino acids, and nucleotide bases.  
 CC Nucleotide sequences, proteins, vectors and host cells from the present  
 CC invention can be used: (a) to identify mosses related to Physcomitrella  
 CC patens; (b) in mapping genomes of mosses related to Physcomitrella  
 CC patens; (c) in the determination of functional TCNRP regions; (f) and in the  
 CC (e) in the determination of functional TCNRP regions; (f) and in the  
 CC cellular production of fine chemicals. AAH44262 to AAH44262 encode the  
 CC Physcomitrella patens TCNRP proteins given in AAB99889 to AAB99889.  
 CC AAH44262 to AAH44262 represent nucleotide sequence used in the  
 CC exemplification of the present invention

XX

XX Sequence 1962 BP; 517 A; 416 C; 518 G; 511 T; 0 U; 0 Other;

XX

## Alignment Scores:

Pred. No.: 8,46e-189 Length: 1962

Score: 1829.50 Matches: 332

Percent Similarity: 83.76% Conservative: 65

Best Local Similarity: 70.04% Mismatch: 74

Query Watch: 72.74% Indels: 3

DB: 4 Gaps: 2

US-10-031-331B-40 (1-473) x AAH44262 (1-1962)  
 QY 1 HistHrValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp 20  
 Db 415 CATCTGTGGAGCCTAGCGTTGAGGCAATGATGCTTGAATTCGACGCGCTCCAAACTCGAT 474  
 QY 21 LysGluGluArgProGluIleLeuSerMetLeuProLeuGluGluLysCysLeuLeu 40  
 Db 475 AAAGAGAACGACCCGAGATTTTGTCTGCTGTTCGCCCATATGTAACACAGATGTCATG 534  
 QY 41 GluLeuGlyAlaGlyIleGlyArgPheThrGlyGluLeuAlaGluLysAlaGlyGlnVal 60  
 Db 535 GAGCTCGGAGCAGGATCGGTGGTGTACTGTGAGCTTGCACAGCATGTCAGGTTCATGTG 594  
 QY 61 IleAlaLeuAspPheIleGluSerAlaIleLysValLeuValIleAsnGlyHisTyr 80  
 Db 595 CTTCGCGATTTTCGAGAGATCTCATCAAGAAGACGAGGATGTGAACGCTCACTAC 654  
 QY 81 LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProHis 100  
 Db 655 AACACATCGATTTCAATGTGCGATGACCTCTCCAGACCTGAATATTCACACAGGT 714  
 QY 101 SerLeuAspValIlePheSerAsnTrpLeuLeuMetTyrLeuSerAspGluValGlu 120  
 Db 715 TCTGCGGATCTCGTGTGTTTCAAAATGCGCTTCTCATGTACTTCTGTGACGAGGTTAAA 774  
 QY 121 AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyGlyTyrIlePhePheArgGlu 140  
 Db 775 GCGTTAGCATCACCGTTATGGAGTGGCTCAGGCGCTGGAGGATACATTTCTTCAGAAA 834  
 QY 141 SerCysPheHisGlnSerGlyAspHisLysArgLysSerAsnProThrHisTyrArgGlu 160  
 Db 835 TCCTGCTTCCACCATCAGGATCAGGATCAACAGCGAAGAACAACTCTACTCCTACCTCAA 894  
 QY 161 ProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSer 180  
 Db 895 CCAACAGGATACAGCAATCTTCCAGCAGCGCTCATATCGAAGAG---GATGGGTCTTAT 951  
 QY 181 TyrGluLeuSerLeuLeuSerCysLysCysIleGlyAlaTyrValArgAsnLysValAsn 200  
 Db 952 TTCAGTTGAATGTGGTGGATGCAATGTGCGACATCATCTGCGAATATGAGAGAAAT 1011  
 QY 201 GlnAsnGlnIleSerTrpLeuTrpGlnLysVal-----AspSerLysAspAspLysGly 218  
 Db 1012 CAAAACCCAGGTGTGTGTTTATGGAGGAAAGTTTCAGTCCGATCGGACCTCGAGAGCGTGT 1071  
 QY 219 PheGlnArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArg 238  
 Db 1072 TTCAGAGAGTTTGGACACCAACAGTACACGTCACCTGGAATCTGCGTTACGAGGT 1131  
 QY 239 ValPheGlyProGlyTyrValSerThrGlyGlyTyrGluThrThrLysGluPheValSer 258  
 Db 1132 ATTTTGGAGAGGAGTTTGTAGCAGCGGTGGATCGAACCACCAAGAGCTTTTGTAGT 1191  
 QY 259 MetLeuAspLeuLysProGlyGlnLysValLeuAspValGlyCysGlyIleGlyGly 278  
 Db 1192 ATGTCTGACTTGAAGCCAGGACAGCGTGTCTTGCATGCGATGCGGATCGGAGGTGT 1251  
 QY 279 AspPheTyrMetAlaGluThrPheAspValGluValGlyPheAspLeuSerValAsn 298  
 Db 1252 GATTTCTACATGCGCGAGAGATATGATGCTGAAGTTGTGGCATCGACCTGCTCTAAAT 1311  
 QY 299 MetIleSerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluVal 318  
 Db 1312 ATGATTTGCTTGTCTTGAACGATCGATCGGAGAAAATGTGCACTGCGAGTTGAAGTT 1371  
 QY 319 AlaAspCysThrLysIleAsnTyrProAspAsnSerPheAspValIleTyrSerArgAsp 338  
 Db 1372 GGGATTTGCCACCAAGATTAAATACCTCAGCATCTTTTGTATGTCATCTACAGTCTGTAT 1431  
 QY 339 ThrIleLeuHisIleGlnAspLysProAlaLeuPheArgSerPheTyrValTrpLeuLys 358  
 Db 1432 ACCATTTCTACATTTCAAGATAAACCTCGGCTTTTTCACGCGTTTATATATGTTGAGAG 1491

```

QY 359 ProGlyGlyLeuValLeuLeuSerAspTyrCysValLeuAlaGlyProProSerProGlu 378
DB 1492 CCGGAGGTCGGGCTTATCAGTACTGATGAGCTCCCAAACTCCGTCGGCGAG 1551
QY 379 PheAlaAlaTyrIleLeuGlnArgGlyTyrAspLeuHisAspValLeuGlyTyrGlyGln 398
DB 1552 TTCGCTGATACATTCACGAGAGGGTTATGATCTCCATAGCGTTCAGAGTACGAGAG 1611
QY 399 MetLeuLeuAspAlaGlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIle 418
DB 1612 ATGCTGGGAAGATGCGGTTTTGTGGAAGTGTGTCAGAGGACCGCACGATCAGTTCAAT 1671
QY 419 ArgValLeuArgLeuGluLeuGluThrValGluLeuGluLeuAspValPheIleSerAsp 438
DB 1672 GAAAGTTTACAGAGGAGCTAGCCACCACTGAGCAGGTCGTGACCACTTCATCAAGCAT 1731
QY 439 PheSerGluGluAspTyrAsnAspIleValGlyGlyTyrAsnAspIleValArgThr 458
DB 1732 TTCCTCGAGGAGGATTAATACATCACTGTGAGCGGATGGAAGATGAGTCAAGCGCTGT 1791
QY 459 AlaLeuGlyGluGlnArgTyrGlyLeuPheValAlaLeuLeuLys 472
DB 1792 TCGAATGACGAACAGAGTGGGACTCTTCATAGCCTACAAAG 1833

RESULT 4
AADI6798
ID AADI6798 standard; DNA; 1120 BP.
XX
AC AADI6798;
XX
DT 29-NOV-2001 (first entry)
XX
DE Spinach PEAMT truncated DNA.
XX
KW Spinach; S-adenosyl-L-methionine, phosphoethanolamine N-methyltransferase;
KW PEAMT; cellular intermediate; phospho-dimethylethanolamine; choline;
KW phosphomono-methylethanolamine; phosphocholine; phosphatidylcholine;
KW glycine betaine; choline-O-sulphate; lipid content alteration;
KW osmotic stress tolerance; nutritional value; transgenic plant;
KW cryoprotectant; ds.
XX
OS Spinacia oleracea.
XX
FH Key Location/Qualifiers
FT CDS 254..1120
FT /tag= a
FT /product= "Spinach PEAMT truncated protein"
FT /note= "CDS does not include stop codon"
FT /partial
XX
FN W0200168870-A2.
XX
PD 20-SEP-2001.
XX
PF 15-MAR-2001; 2001WO-US008352.
XX
PR 15-MAR-2000; 2000US-00525885.
XX
PA (UYFL) UNIV FLORIDA.
XX
PA (UYCA-) UNIV CARNEGIE MELLON.
XX
PI Hanson AD, Nuccio ML, Henry SA;
XX
PI WPI; 2001-565796/63.
XX
DR P-PSDB; AAB09761.
XX
DR
XX
PT New plant S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase
PT polypeptides, useful for modulating the levels of cellular intermediates
PT such as-phosphedimethylethanolamine and for altering the lipid content in
PT plants cells.
XX
PS Claim 18; Page 109; 158pp; English.

```

XX

The present sequence is spinach S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase (PEAMT) truncated DNA. The PEAMT sequences are useful for modulating the levels of cellular intermediates such as phosphodimethylethanolamine, phosphomono-methylethanolamine, choline, phosphocholine, phosphatidylcholine, choline-O-sulphate or glycine betaine. They are useful for altering the lipid content in plant cells. The polynucleotides are also useful for improving the osmotic stress tolerance of a plant and increasing the cryoprotectant properties of a plant. The present invention also relates to methods and compositions comprising PEAMT used for generating transgenic plants with increased nutritional value

XX SQ Sequence 1120 BP; 304 A; 218 C; 256 G; 342 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 3,41e-128 Length: 1120  
Score: 1270.00 Matches: 235  
Percent Similarity: 94.34% Conservative: 15  
Best Local Similarity: 88.68% Mismatches: 15  
Query Match: 50.50% Indels: 0  
DB: 5 Gaps: 0

US-10-031-331b-40 (1-473) x AADI6798 (1-1120)

```

QY 1 HisThrValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp 20
DB 317 CACTCTGTTGATTTGACTGTTGAGGCTATGATGCTTGAATTCACAGCTTCAGATCTTGAC 376
QY 21 LysGluGluArgProGluIleLeuSerMetLeuProProLeuGluGlyLysCysLeuLeu 40
DB 377 AAGTGGAGCGACCTGAGGTACTTTCATGCTTCCACCTTATGAGGAAAGTCTGCTTA 436
QY 41 GluLeuGlyAlaGlyIleGlyArgPheThrGlyGluLeuAlaGlyValGlnVal 60
DB 437 GAACCTGGTGCTGATGTTGCTGCTGTTTACTGGTGAATTCGCCGAAAGCTAGCAGGTC 496
QY 61 IleAlaLeuAspPheIleGluSerAlaIleLysAsnGluValIleAsnGlyHisTyr 80
DB 497 ATCGCTCTGGATTTCAATGAGAGTGTATTAAGAAGATGAAGAATTAATGGGCATTAC 556
QY 81 LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProHis 100
DB 557 ABAATGTGAAGTTTATGTTGCTGATGTGACATCTCCAGTCTCAACATTCACCAAT 616
QY 101 SerLeuAspValIlePheSerAsnTrpLeuLeuMetTyrLeuSerAspGluValGlu 120
DB 617 TCCGTGGATATCATATTCCTCAATTCGCTACTCATGTATCTTTCTGATGAAGAGGTTGAG 676
QY 121 AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyTyrIlePhePheArgGlu 140
DB 677 CGTCTGGTTGAAGGATGTTGAATGTTGAAGCCAGGAGGATACATTTCTTCAGAGAA 736
QY 141 SerCysPheHisGlnSerGlyAspHisLysArgLysSerAsnProThrHisTyrArgGlu 160
DB 737 TCTTGTGTTTTCATCAATCAGGAGATCAAGCGCAAGCAATCCCAACCTACCTACCGTGA 796
QY 161 ProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSer 180
DB 797 CCTAGGTTCTACACCAAGATCTTCAAGAATGCCATATGCAAGATGATTTCTGGGAATCC 856
QY 181 TyrGluLeuSerLeuLeuSerCysLysGlyIleGlyValTyrValArgAsnLysAsn 200
DB 857 TATGAGCTCTCCCTTCCTCAATTCGCTCAATGATGAGCTTATGTCANAGCAAGAGAT 916
QY 201 GlnAsnGlnIleSerTrpLeuTrpGlnLysValAspSerLysAspAspLysGlyPheGln 220
DB 917 CAGAACCAAGATAGCTGTTATGCGCAGAAAGTGTGATTCAGAGATGACACAGGGGTTCCAG 976
QY 221 ArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArgValPhe 240
DB 977 CGATCTCTGGATCTAGTCAATCAAGATTTACAGCATCTACTCGTTATGAGCGGTATTT 1036

```









Db 514 CCAAGGTTTACACTAAGGTAATTTAAAGAGGGCCGATGCAATTCATCAGAGTGCGAGCTCC 573

Qy 181 TyrGluLeuSerLeuLeuSerCysAlaValGluPheGluValAlaAaspCysThrLys 323  
 Db 713 CTGAGCATGCTATTTGGACGCAAGTGCGAGTTCGAGTTGCTGATTGACACCG 654

Qy 324 IleAsnTyrProAspAsnSerPheAspValIleTyrSerArgAspThrIleLeuHisIle 343  
 Db 653 AAGACATACCCAGACAAATACATTTGATGTATCTACAGCCGTGACACCATCTTCACATA 594

Qy 344 GlnAspLysProAlaLeuPheArgSerPheTyrIleTyrTrpLeuLysProGlyGlyLysVal 363  
 Db 593 CAAAGTAAACCCGCTTTGTTTAGAAGTTCTTCAATGGCTAAACCTGCTGTAAGTTC 534

Qy 364 LeuIleSerAspTyrCysLysLysAlaGlyProSerProGluPheAlaLysVal 383  
 Db 533 CTAATCAGCGATTAATCTAGGAGTCCAGAAACCATCTGACAGGTTTGGCTATACATT 474

Qy 384 LysGlnArgGlyTyrAspLeuHisAspValLysGluTyrGlyGlnMetLeuLysAspAla 403  
 Db 473 AAGCAGAGGGGTACGACCTTCATGATGTAGAGACTTATGACAGATGCTAGAGAATGCT 414

Qy 404 GlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIleArgValLeuArgLys 423  
 Db 413 GGTTCATGATGTCATTTGCCAGAACCCGACGATCAGTCTCTGAAAGTTTACAGAGG 354

Qy 424 GluLeuGluThrValGluLysGluLysAspValPheIleSerAspPheSerGluGluAsp 443  
 Db 353 GAGCTAGCCGAAGTTGAAAGAACCAAGATGATTTTCTAGCCGCTTTGGTCAGAGGAGC 294

Qy 444 TyrAsnAspIleValGlyGlyTyrAsnAspLysLeuArgArgThrAlaLysGlyGluGln 463  
 Db 293 TATGACGATATTGCTGCTGGTGGAACCGCAAACTTCATAGAGCTCTGCTGCTGAGCAG 234

Qy 464 ArgTrpGlyLeuPheValAlaLysLys 472  
 Db 233 AGGTGGGGGCTGTTTCATTTGGGACCAAG 207

RESULT 10  
 ACL18347/c  
 ID ACL18347 standard; DNA; 575 BP.  
 XX  
 AC ACL18347;  
 XX  
 DT 27-OCT-2003 (revised)  
 DT 17-OCT-2003 (first entry)  
 XX  
 DE DNA clone originating in barley containing SNP encoding sequence #8338.  
 XX  
 KW Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;  
 KW gene; ss.  
 XX  
 OS Hordeum vulgare; var. (cul.Haruna Nijo).  
 XX  
 PN WO2003057877-A1.  
 XX  
 PD 17-JUL-2003.  
 XX  
 PF 16-DEC-2002; 2002WO-IB005403.  
 XX  
 PR 20-DEC-2001; 2001JP-00387059.  
 PR 20-DEC-2001; 2001JP-00387131.  
 PR 20-DEC-2001; 2001JP-00403299.  
 PR 20-DEC-2001; 2001JP-00403300.  
 XX  
 XX 27-SEP-2002; 2002JP-00327515.  
 PA (UTNI-) UNIV JAPAN OKAYAMA.  
 XX  
 PI Sato K, Takeda K, Kohara Y;  
 XX  
 DR WPI; 2003-587127/55.  
 XX  
 PT Single nucleotide polymorphism sites in barley varieties and DNA  
 PT sequences containing them for analysis and identification of barley

Disclosure; SEQ ID XX; 284pp; Japanese.

The present invention relates to oligonucleotide clones originating in barley (Hordeum vulgare) which contain single nucleotide polymorphisms (SNP). The oligonucleotides may be used for analysis of SNPs among barley varieties, identification of particular varieties and creation of new varieties by analysis, isolation of specific genes and production of new varieties by transformation of barley varieties with them and production of new barley varieties with desired properties. The present sequence represents an oligonucleotide clone DNA sequence featured in the specification. The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published-pct-sequences

Sequence 777 BP; 230 A; 188 C; 146 G; 212 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	7-73e-74	Length:	777
Score:	769.00	Matches:	139
Percent Similarity:	86.77%	Conservative:	25
Best Local Similarity:	73.54%	Mismatches:	25
Query Match:	30.58%	Indels:	0
DB:	8	Gaps:	0

US-10-031-331b-40 (1-473) x ACL19405 (1-777)

Qy 284 GluThrPheAspValGluValGlyPheAspLeuSerValAsnMetIleSerPheAla 303  
 Db 773 GAAACATATGATGTCATGTTGTCATGATGCTTCATCAATACATGTTTCATTGCA 714

Qy 304 LeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluValAlaAaspCysThrLys 323  
 Db 713 CTGAGCATGCTATTTGGACGCAAGTGCGAGTTCGAGTTGCTGATTGACACCG 654

Qy 324 IleAsnTyrProAspAsnSerPheAspValIleTyrSerArgAspThrIleLeuHisIle 343  
 Db 653 AAGACATACCCAGACAAATACATTTGATGTATCTACAGCCGTGACACCATCTTCACATA 594

Qy 344 GlnAspLysProAlaLeuPheArgSerPheTyrIleTyrTrpLeuLysProGlyGlyLysVal 363  
 Db 593 CAAAGTAAACCCGCTTTGTTTAGAAGTTCTTCAATGGCTAAACCTGCTGTAAGTTC 534

Qy 364 LeuIleSerAspTyrCysLysLysAlaGlyProSerProGluPheAlaLysVal 383  
 Db 533 CTAATCAGCGATTAATCTAGGAGTCCAGAAACCATCTGACAGGTTTGGCTATACATT 474

Qy 384 LysGlnArgGlyTyrAspLeuHisAspValLysGluTyrGlyGlnMetLeuLysAspAla 403  
 Db 473 AAGCAGAGGGGTACGACCTTCATGATGTAGAGACTTATGACAGATGCTAGAGAATGCT 414

Qy 404 GlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIleArgValLeuArgLys 423  
 Db 413 GGTTCATGATGTCATTTGCCAGAACCCGACGATCAGTCTCTGAAAGTTTACAGAGG 354

Qy 424 GluLeuGluThrValGluLysGluLysAspValPheIleSerAspPheSerGluGluAsp 443  
 Db 353 GAGCTAGCCGAAGTTGAAAGAACCAAGATGATTTTCTAGCCGCTTTGGTCAGAGGAGC 294

Qy 444 TyrAsnAspIleValGlyGlyTyrAsnAspLysLeuArgArgThrAlaLysGlyGluGln 463  
 Db 293 TATGACGATATTGCTGCTGGTGGAACCGCAAACTTCATAGAGCTCTGCTGCTGAGCAG 234

Qy 464 ArgTrpGlyLeuPheValAlaLysLys 472  
 Db 233 AGGTGGGGGCTGTTTCATTTGGGACCAAG 207

RESULT 10  
 ACL18347/c  
 ID ACL18347 standard; DNA; 575 BP.  
 XX  
 AC ACL18347;  
 XX  
 DT 27-OCT-2003 (revised)  
 DT 17-OCT-2003 (first entry)  
 XX  
 DE DNA clone originating in barley containing SNP encoding sequence #8338.  
 XX  
 KW Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;  
 KW gene; ss.  
 XX  
 OS Hordeum vulgare; var. (cul.Haruna Nijo).  
 XX  
 PN WO2003057877-A1.  
 XX  
 PD 17-JUL-2003.  
 XX  
 PF 16-DEC-2002; 2002WO-IB005403.  
 XX  
 PR 20-DEC-2001; 2001JP-00387059.  
 PR 20-DEC-2001; 2001JP-00387131.  
 PR 20-DEC-2001; 2001JP-00403299.  
 PR 20-DEC-2001; 2001JP-00403300.  
 XX  
 XX 27-SEP-2002; 2002JP-00327515.  
 PA (UTNI-) UNIV JAPAN OKAYAMA.  
 XX  
 PI Sato K, Takeda K, Kohara Y;  
 XX  
 DR WPI; 2003-587127/55.  
 XX  
 PT Single nucleotide polymorphism sites in barley varieties and DNA  
 PT sequences containing them for analysis and identification of barley

PT varieties and production of barley transformants with desired  
PT characteristics.  
PS Disclosure; SEQ ID XX; 284pp; Japanese.  
XX The present invention relates to oligonucleotide clones originating in  
CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms  
CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley  
CC varieties, identification of particular varieties and genotype-phenotype  
CC analysis, isolation of specific genes and creation of new varieties by  
CC transformation of barley varieties with them and production of new barley  
CC varieties with desired properties. The present sequence represents an  
CC oligonucleotide clone DNA sequence featured in the specification. The  
CC sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to  
CC standardise OS field)  
XX SQ Sequence 575 BP; 128 A; 160 C; 127 G; 159 T; 0 U; 1 Other;

Alignment Scores:  
Pred. No.: 7.28e-74 Length: 575  
Score: 767.50 Matches: 138  
Percent Similarity: 83.07% Conservative: 19  
Best Local Similarity: 73.02% Mismatches: 31  
Query Match: 30.52% Indels: 1  
DB: 8 Gaps: 1

US-10-031-331B-40 (1-473) x ACL18347 (1-575)  
QY 78 GlyH1g---TyrLysAsnVallypHeMetCysAlaAspValThrSerProThrLeuSer 96  
DB 567 GGGCAGCATCCAGCAGCAGCATCCCTTCATGTCGCGGAGTGCAGTGCAGGCTCAAG 508  
QY 97 PheProPheHisSerLeuAspValIlePheSerAsnTrpLeuMetTrpLeuSerAsp 116  
DB 507 ATCAGAGCAACTCCATCGATCGATCGCTTCCTCACTGCTGCTGCTGCTGCTGCTGCT 448  
QY 117 GluGluValGluAsnLeuValGluArgMetLeuLysTrpLeuLysProGlyTyrIle 136  
DB 447 GAGGAGTTGAGAGCTGATGGCAGATAGTGAAGTGGCTGAAGCTGCTGCTGCTGCTGCT 388  
QY 137 PhePheArgGluSerCysPheHisGlnSerGlyAspHisIlyAsArgLysSerAsnProThr 156  
DB 387 TTTATCGAGGATTCGCTTCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 328  
QY 157 HisTyrArgGluProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGly 176  
DB 327 CACTACCGGAGCGAGGTTTACACCAAGGTGTTCAAGGATGCCACTCCTATGACCAA 268  
QY 177 SerGlyAsnSerTyrGluLeuSerLeuLeuSerCysLysCysAlleGlyAlaTyrValArg 196  
DB 267 GAGGGGAACTCTTTGAGCTTTCTGCTGTAACCTCCAGTGTATGAGGCTTATGTGAAA 208  
QY 197 AsnLysLysAsnGlnAsnGlnIleSerTrpLeuTrpGlnLysValAspSerLysAspAsp 216  
DB 207 AGCAAGAGAACCAAGCAAGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 148  
QY 217 LysGlyPheGlnArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyr 236  
DB 147 AAGGCTTTCAGAGATTCCTGCAATATGTCAGTAATCAATCCATCCATGGAATCTTCGCTAC 89  
QY 237 GluArgValPheGlyProGlyTyrValSerThrGlyTyrGluThrThrLysGluPhe 256  
DB 87 GAGGCGCTGTTTGGGAGGCTTATGTTAGCACCGGTGATTCGAGGAGGATTCGAGGAGATTT 28  
QY 257 ValSerMetLeuAspLeuLysProGly 265  
DB 27 GTGACCAAGCTGGACCTGGAAGCTGTC 1  
RESULT 11  
ACL19398/c  
ID ACL19398 standard; DNA; 566 BP.

XX ACL19398;  
AC 17-OCT-2003 (first entry)  
DT DNA clone originating in barley containing SNP encoding sequence #9389.  
DE Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;  
KW Gene; ss.  
XX Hordeum vulgare.  
OS WO2003057877-A1.  
PN 17-JUL-2003.  
PD 16-DEC-2002; 2002WO-IB005403.  
PP 20-DEC-2001; 2001JP-00387059.  
PR 20-DEC-2001; 2001JP-00387131.  
PR 20-DEC-2001; 2001JP-00403299.  
PR 20-DEC-2001; 2001JP-00403300.  
PR 27-SEP-2002; 2002JP-00327515.  
XX (UTNI-) UNIV JAPAN OKAYAMA.  
XX Sato K, Takeda K, Kohara Y;  
XX WPI; 2003-587127/55.  
XX Single nucleotide polymorphism sites in barley varieties and DNA  
PT sequences containing them for analysis and identification of barley  
PT varieties and production of barley transformants with desired  
PT characteristics.  
XX Disclosure; SEQ ID XX; 284pp; Japanese.  
XX The present invention relates to oligonucleotide clones originating in  
CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms  
CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley  
CC varieties, identification of particular varieties and genotype-phenotype  
CC analysis, isolation of specific genes and creation of new varieties by  
CC transformation of barley varieties with them and production of new barley  
CC varieties with desired properties. The present sequence represents an  
CC oligonucleotide clone DNA sequence featured in the specification. The  
CC sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published-pct-sequences  
XX SQ Sequence 566 BP; 168 A; 140 C; 97 G; 156 T; 0 U; 5 Other;

Alignment Scores:  
Pred. No.: 4.47e-72 Length: 566  
Score: 751.00 Matches: 143  
Percent Similarity: 84.57% Conservative: 16  
Best Local Similarity: 76.06% Mismatches: 28  
Query Match: 29.86% Indels: 1  
DB: 8 Gaps: 0

US-10-031-331B-40 (1-473) x ACL19398 (1-566)  
QY 163 PheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSerTyrGlu 182  
DB 564 TTTTACACTANGATTATTAANGNGCCATGCCATGATGATGATGATGATGATGATGATGAT 505  
QY 183 LeuSerLeuLeuSerCysLysCysAlleGlyAlaTyrValArgAsnLysLysAsnGlnAsn 202  
DB 504 CTTTCTCTGCTTACTTGCAGTGTGTGTGGAGCTTATGTGAGAGACACAGAGATCAAAAC 445  
QY 203 Gln-IleSerTrpLeuTyrGlnLysValAspSerLysAspAspLysGlyPheGlnArgph 222  
DB 444 CAGNATATGTTGCTATGTCGCAAAAGTCAACTCAACAGAGATCGCGGCTTCAAGATT 385

QY 222 cLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArgValPheGlyPr 242  
 DB 384 TTGGATATGTCAGTACAAACCAAGTGAATATTACGCTATGAGGTGTTTGGCA 325  
 QY 242 cGlyTyrValSerThrGlyGlyTyrGluThrThrValSerMetLeuAspLe 262  
 DB 324 AGTTTGTGAGCAGCTGGTGAATGAGACTACAAAGAAATTTGGACTTCTGGATCT 265  
 QY 262 wlyserGlyGlnPheValLeuAspValGlyCysGlyIleGlyGlyGlyAspPheTyrMe 282  
 DB 264 TAAACCTGGCAGAGGTGCTGATGTTGATGTAATCGGGGTGGATTTTAT 205  
 QY 282 talGluThrPheAspValGluValValGlyPheAspLeuSerValAsnMetIleSerPh 302  
 DB 204 GCGCGAAACTATGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTT 145  
 QY 302 calaleuGluArgSerIleGlyLeuLysCysValValGluPheGluValAlaAspCysTh 322  
 DB 144 TGCATCGAGCATCTATTGGACGCAAGTCGCGAGTTCGAGTTTGATGTTGATGTTG 85  
 QY 322 rlylIleAsnTyrProPheAspSerPheAspValIleTyrSerArgAspThrIleLeuHi 342  
 DB 84 CACGAGACATACCCAGACATATCTTATGATGTTATCTACAGCGGTGACACCATCTTCA 25  
 QY 342 sileGlnAspLysProAlaLeu 349  
 DB 24 CATAACAGATAAACCGCTTGG 3  
 RESULT 12  
 ACL18342/c  
 ID ACL18342 standard; DNA; 650 BP.  
 XX  
 AC ACL18342;  
 XX  
 27-OCT-2003 (revised)  
 DT 17-OCT-2003 (first entry)  
 XX  
 DE DNA clone originating in barley containing SNP encoding sequence #8333.  
 XX  
 KW Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;  
 XX Gene; ss.  
 OS Hordeum vulgare; var. (cul.Haruna Nijo).  
 XX  
 PN WO2003057877-A1.  
 XX  
 PD 17-JUL-2003.  
 XX  
 PF 16-DEC-2002; 2002WO-1B005403.  
 XX  
 PR 20-DEC-2001; 2001JP-00387059.  
 PR 20-DEC-2001; 2001JP-00387131.  
 PR 20-DEC-2001; 2001JP-00403299.  
 PR 20-DEC-2001; 2001JP-00403300.  
 PR 27-SEP-2002; 2002JP-00327515.  
 XX  
 PA (UYN1-) UNIV JAPAN OKAYAMA.  
 XX  
 PI Sato K, Takeda K, Kohara Y;  
 XX  
 DR WPI; 2003-587127/55.  
 XX  
 PT Single nucleotide polymorphism sites in barley varieties and DNA  
 PT sequences containing them for analysis and identification of barley  
 PT varieties and production of barley transformants with desired  
 PT characteristics.  
 XX  
 PS Disclosure; SEQ ID XX; 284pp; Japanese.  
 XX  
 CC The present invention relates to oligonucleotide clones originating in  
 CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms  
 CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley  
 CC varieties, identification of particular varieties and genotype-phenotype

CC analysis, isolation of specific genes and creation of new varieties by  
 CC transformation of barley varieties with them and production of new barley  
 CC varieties with desired properties. The present sequence represents an  
 CC oligonucleotide clone DNA sequence featured in the specification. The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to  
 CC standardise OS field)  
 XX  
 SQ Sequence 650 BP; 123 A; 194 C; 177 G; 156 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 1,31e-71 Length: 650  
 Score: 747.50 Matches: 137  
 Percent Similarity: 86.17% Conservative: 25  
 Best Local Similarity: 72.87% Mismatches: 25  
 Query Match: 29.72% Indels: 1  
 DB: 8 Gaps: 1

US-10-031-331B-40 (1-473) x ACL18342 (1-650)

QY 1 KieThrValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp 20  
 DB 564 CACTCCAGGACCTCACCGTCCAGTCCATGCTCGACTCCGCGCCAGGACCTCCAC 505  
 QY 21 LysGluGluArgProGluIleLeuSerMetLeuProLeuGluGlyLysCysLeuLeu 40  
 DB 504 AAGGAGGAGCGCCAGAGGTGCTGGCCATCTCTATGAGGGCAAGACTGTGCTG 445  
 QY 41 GluLeuGlyAlaGlyIleGlyArgPheThrGlyGluLeuAlaGluLysAlaGlyGlnVal 60  
 DB 444 GAGCTCGCGCGCATCGCGCTTTACCGCGAGCTGGCCAGAGGCGCGCCACGTC 385  
 QY 61 IlealLeuAspPheIleGluSerAlaIleLysLysAsnGluValIleAsnGlyHis--- 79  
 DB 384 ATCCGCTGGACTTCATCGACGCTCATCAGAGAAAGAGAGAGATCAATGGGCGCATC 325  
 QY 80 TyrLysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProPro 99  
 DB 324 CACAGCAACATACCTTCATGTCGCGCGAGTGAAGTCCGCGAGCTCAAGATCGAGGAC 265  
 QY 100 HisSerLeuAspValIlePheSerAsnTyrLeuLeuMetTyrLeuSerAspGluVal 119  
 DB 264 AACTCCATCGACATCGCTTCTCCACTGCTGCTCATGTACTCATCGACGAGAGGT 205  
 QY 120 GluAsnLeuValGluArgMetLeuLysTrpLeuLysProGlyTyrIlePhePheArg 139  
 DB 204 GAGAAGCTGATTGGCAGATAGTGAAGTGGCTGAAAGCTGGTGGACATATTTTATCAG 145  
 QY 140 GluSerCysPheHisGlnSerGlyAspHisLysArgLysSerAsnProThrHisTyrArg 159  
 DB 144 GAATCGTGTTCATCATCAGTCTGCTGATTCCAGAGAGAAAGTGAACCCGACCTACCG 85  
 QY 160 GluProArgPheTyrThrIleAlaPheLysGlyCysHisLeuGlnAspGlySerGlyAsn 179  
 DB 84 GAGCGAGGTTTTACACCAAGGTTCAGAGGATGCCACTCTATGACCAAGAGGGGAAAC 25  
 QY 180 SerTyrGluLeuSerLeuLeuSer 187  
 DB 24 TCTTTGAGCTTCTCTGTGTAAT 1

## RESULT 13

ACL18325/c

ID ACL18325 standard; DNA; 622 BP.

XX

AC ACL18325;

XX

27-OCT-2003 (revised)

DT 17-OCT-2003 (first entry)

XX

DE DNA clone originating in barley containing SNP encoding sequence #8316.

XX

KW Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;

KW gene; ss.  
 XX Hordeum vulgare; esp. spontaneum.  
 XX WO2003057877-A1.  
 PN 17-JUL-2003.  
 PD 16-DEC-2002; 2002WO-IB005403.  
 PP 20-DEC-2001; 2001JP-00387059.  
 PR 20-DEC-2001; 2001JP-00387131.  
 PR 20-DEC-2001; 2001JP-00403299.  
 PR 20-DEC-2001; 2001JP-00403300.  
 PR 27-SEP-2002; 2002JP-00327515.  
 XX (UYNI-) UNIV JAPAN OKAYAMA.  
 XX Sato K, Takeda K, Kohara Y;  
 PI WPI; 2003-587127/55.  
 DR  
 XX Single nucleotide polymorphism sites in barley varieties and DNA  
 PT sequences containing them for analysis and identification of barley  
 PT varieties and production of barley transformants with desired  
 PT characteristics.  
 XX  
 XX Disclosure; SEQ ID XX; 284pp; Japanese.  
 XX  
 CC The present invention relates to oligonucleotide clones originating in  
 CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms  
 CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley  
 CC varieties, identification of particular varieties and genotype-phenotype  
 CC analysis, isolation of specific genes and creation of new varieties by  
 CC transformation of barley varieties with them and production of new barley  
 CC varieties with desired properties. The present sequence represents an  
 CC oligonucleotide clone DNA sequence featured in the specification. The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at [ftp.wipo.int/pub/published-pct-sequences](http://ftp.wipo.int/pub/published-pct-sequences). (Updated on 27-OCT-2003 to  
 CC standardise OS field)  
 XX  
 SQ Sequence 622 BP; 118 A; 187 C; 173 G; 144 T; 0 U; 0 Other;  
 XX  
 Alignment Scores:  
 Pred. No.: 6.02e-70 Length: 622  
 Score: 732.00 Matches: 135  
 Percent Similarity: 87.10% Conservative: 27  
 Best Local Similarity: 72.58% Mismatches: 24  
 Query Match: 29.11% Indels: 1  
 DB: 8 Gaps: 0  
 US-10-031-331b-40 (1-473) x ACL18325 (1-622)  
 QY 287 AspValGluValValGlyPheAspLeuSerValAsnMetIleSerPheAlaLeuGluArg 306  
 .db 619 GATGTCATGCTCTGGCATGATCTTTCTC-AACATGTTTCTTCTCGCATCGAGGCT 561  
 QY 307 SerIleGlyLeuLysCysAlaValGluPheGluValAlaAspCysThrLysIleAsnTyr 326  
 .db 560 GCATCGGTGCTCTGCTGCTGGTTGAGTTGAGTTGCTGCTGACACCCAGGATAC 501  
 QY 327 ProAspAsnSerPheAspValIleTyrSerArgAspThrIleLeuHisIleGlnAspLys 346  
 .db 500 CCAGAGMACATGCTGATGTCTATACAGCGGTGACACCATCTCCACATTCAGCAGAG 441  
 QY 347 ProAlaLeuPheArgSerPheTyrIleTyrPheLysProGlyLysValLeuIleSer 366  
 .db 440 CTTGCTCTGTTCCAGAACTTCTTCAAGTGGCTGAAGCTTGGTGGCAAGTGTCTGATCA 381  
 QY 367 AspTyrCysLysIleAlaGlyProPheSerProGluPheAlaIleTyrIleLysGlnArg 386  
 .db 380 GACTACTGACGAGGCGCTGGGACGCCATCAGAGGATTCCTGCTGATCATCAGCAGAGA 321

QY 387 GlyTyrAspLeuHisAspValLysGluTyrGlyGlnMetLeuLysAspAlaGlyPheVal 406  
 .db 320 GGCTATGACCTCCATGATGTGAAGACCTACGGAAGATGCTTGAGGATGCGGGTTCCAC 261  
 QY 407 AspValLeuAlaGluAspArgThrGluGlnPheIleArgValLeuArgLysGluLeuGlu 426  
 .db 260 GACGTTCATCGCGAGACCCGACCGACCACTTCTCGGGTCTCTGGAGGAGGAGCTGGGC 201  
 QY 427 ThrValGluLysGluLysAspValPheIleSerAspPheSerGluLysAspTyrAsnAsp 446  
 .db 200 GAGACCGAGAGAGAACAGAGGAGGCTTCTCGCGACTTCGCGAGGAGGACTACGACAC 241  
 QY 447 IleValGlyGlyTyrAsnAspLysLeuArgArgThrAlaLysGlyGluGlnArgTyrGly 466  
 .db 140 ATCGTCAACGGTGGAGCGCGAGCTGAAGCGGAGCTCCGCGCGAGCAGAGTGGGG 81  
 QY 467 LeuPheValAlaLysLys 472  
 .db 80 CTGTTTCATCGGACCAAG 63  
 RESULT 14  
 ADEB2005  
 ID ADEB2005 standard; cDNA; 487 BP.  
 XX  
 AC ADEB2005;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Arabidopsis thaliana expressed polynucleotide seq id 776.  
 KW genetically modified organism; transgenic organism; plant;  
 KW inhibitor testing; activator testing; modifier testing; fungicide;  
 KW insecticide; genetic function; genetic regulation; cellular metabolism;  
 KW gene; ss.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN US2003115639-A1.  
 XX  
 PD 19-JUN-2003.  
 XX  
 PF 26-JAN-2001; 2001US-00770961.  
 XX  
 PR 27-JAN-2000; 2000US-0178466P.  
 XX  
 PA (GORL/) GORLACH J.  
 PA (ANY/) AN Y.  
 PA (HAMI/) HAMILTON C M.  
 PA (PRIC/) PRICE J L.  
 PA (RAIN/) RAINES T M.  
 PA (YUY/) YU Y.  
 PA (RAME/) RAMEKA J G.  
 PA (PAGE/) PAGE A.  
 PA (MATH/) MATHEN A V.  
 PA (LEDF/) LEDFORD B L.  
 PA (WOES/) WOESSNER J P.  
 PA (HAAS/) HAAS W D.  
 PA (GARC/) GARCIA C A.  
 PA (KRIC/) KRICKER M.  
 PA (SLAT/) SLATER T.  
 PA (DAVI/) DAVIS K R.  
 PA (ALLE/) ALLEN K.  
 PA (HOFF/) HOFFMAN N.  
 PA (HURB/) HURBAN P.  
 XX  
 PI Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;  
 PI Rameka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;  
 PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;  
 PI Hurban P;  
 XX  
 WPI; 2003-810930/76.



PT Novel Arabidopsis thaliana nucleic acids useful for generating  
PT Genetically modified transgenic organisms, for screening biologically  
XX active agents such as fungicides, insecticides.

PS Claim 1; SEQ ID NO 776; 44pp; English.

XX The invention describes a nucleic acid (I) comprising a sequence capable  
CC of hybridising under stringent conditions to any one of 999 fully defined  
CC Arabidopsis thaliana sequences (I) as given in specification e.g., 360,  
CC 1137, 455, 219, 472 nucleotides, etc., or its fragment. (I) is useful as a  
CC hybridisation probe to complementary molecules in a cDNA library. (I) is  
CC also useful for generating genetically modified and transgenic organisms,  
CC usually plant cells and plants. A protein encoded by (I) is useful in  
CC screening assays to determine the effect of candidate inhibitors,  
CC activators or modifiers of the gene product. The protein is also useful  
CC for screening biologically active agents e.g., fungicides and  
CC insecticides. A genetically modified cell, comprising an exogenous  
CC nucleic acid, where the nucleic acid comprises transcription regulatory  
CC sequences operably linked to a sequence capable of hybridising under  
CC stringent conditions to (I) is useful in the study of genetic function  
CC and regulation, for alteration of the cellular metabolism and for  
CC screening compounds that may affect the biological function of the gene  
CC or gene product. This sequence represents an Arabidopsis thaliana  
CC polynucleotide of the invention.

XX Sequence 487 BP; 129 A; 88 C; 131 G; 139 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.: 1,18-68 Length: 487  
Score: 719.00 Matches: 133  
Percent Similarity: 89.51% Conservative: 12  
Best Local Similarity: 82.10% Mismatches: 17  
Query Match: 28.59% Indels: 0  
DB: 9 Gaps: 0

US-10-031-331B-40 (1-473) x ADB82005 (1-487)

QY 182 GluLeuSerLeuLeuSerCysLysCysLeuGlyAlaTyrValArgAnLysLysAenGln 201  
DB 2 GAGCTCTCTATGTTGGTGGCTGCAATGCAATGGGGCTTATGTGAGAACCAAGAATCAG 61  
QY 202 AenGlnIleSerTyrLeuTyrGlnValValSerLysAspLysGlyPheGlnArg 221  
DB 62 AATCAGATTTCCTGGATATGCAAAAAGTCAGCGTCGAGATGACAGGATTTCCAGCGT 121  
QY 222 PheLeuAspThrSerGlnTyrLysCysAenSerIleLeuArgTyrGluArgValPheGly 241  
DB 122 TCTTGGACATGTTCAATACAGCTAGTGGGATCTTCGGCTATGAGCGGTCTTTGGG 181  
QY 242 ProGlyTyrValSerThrGlyTyrGlnThrLysGluPheValSerMetLeuAsp 261  
DB 182 GAGGATATGTGACACTGCTGGATTTGAGACAACTAAAGATTTGTGCGAAGATGAC 241  
QY 262 LeuLysProGlyGlnLysValLeuAspValGlyCysGlyIleGlyGlyLysAspPheTyr 281  
DB 242 CTTAAACCGGACAGAAAGTCTTAGATGTTGGTGTGATCGGTGGAGGTGACTTCTAC 301  
QY 282 MetAlaGluThrPheAspValGluValValGlyPheAspLeuSerValMetIleSer 301  
DB 302 ATGCTTGAGATTTCGATGTTTCATGTTTGGATTCGATCTGCTCGGTCAACATGATCTCT 361  
QY 302 PheAlaLeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluValAlaAspCys 321  
DB 362 TTGCACTGGAGCGGCGCATTTGGACTCAAAATGCTCAGTCTGAGTTTGAAGTCTGATTCG 421  
QY 322 ThrLysIleAenTyrProAspAenSerPheAspValIleTyrSerArgAspThrIleLeu 341  
DB 422 ACCACCAAAACATATCCGATATTTCTTTGATGTCAATTTACAGCCGTGACACTTCTG 481  
QY 342\_Histle\_343  
DB 482 CACATC 487

#### RESULT 15

ACL25708  
ID ACL25708 standard; DNA; 589 BP.

XX ACL25708,

XX 27-OCT-2003 (revised)

DT 17-OCT-2003 (first entry)

XX DNA clone originating in barley containing SNP encoding sequence #15699.

XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;  
XX gene, ss.

XX Hordeum vulgare; var. (cul.Haruna Nijo).

XX WO2003057877-A1.

XX 17-JUL-2003.

XX 16-DEC-2002; 2002WO-IB005403.

XX 20-DEC-2001; 2001JP-00387059.

XX 20-DEC-2001; 2001JP-00387131.

XX 20-DEC-2001; 2001JP-00403299.

XX 20-DEC-2001; 2001JP-00403300.

XX 27-SEP-2002; 2002JP-00327515.

XX (UYMI-) UNIV JAPAN OKAYAMA.

XX Sato K, Takeda K, Kohara Y;

XX WPI; 2003-587127/55.

XX Single nucleotide polymorphism sites in barley varieties and DNA  
XX sequences containing them for analysis and identification of barley  
XX varieties and production of barley transformants with desired  
XX characteristics.

XX Disclosure; SEQ ID XX; 284pp; Japanese.

XX The present invention relates to oligonucleotide clones originating in  
XX barley (Hordeum vulgare), which contain single nucleotide polymorphisms  
XX (SNP). The oligonucleotides may be used for analysis of SNPs among barley  
XX varieties, identification of particular varieties and genotype-phenotype  
XX analysis, isolation of specific genes and creation of new varieties by  
XX transformation of barley varieties with them and production of new barley  
XX varieties with desired properties. The present sequence represents an  
XX oligonucleotide clone DNA sequence featured in the specification. The  
XX sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to  
XX standardise OS field)

XX Sequence 589 BP; 161 A; 120 C; 166 G; 142 T; 0 U; 0 Other;

XX Alignment Scores:

Pred. No.: 3,62e-66 Length: 589

Score: 697.00 Matches: 131

Percent Similarity: 87.95% Conservative: 15

Best Local Similarity: 78.92% Mismatches: 20

Query Match: 27.71% Indels: 0

DB: 8 Gaps: 0

US-10-031-331B-40 (1-473) x ACL25708 (1-589)

QY 1 HisThrValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp 20

DB 89 CACTCCAGGACCTCACTGTGTGAGGCATGAGTCTCGACTTCGCTGCCCGCATCTGAC 148

QY 21 LysGluGluArgProGluIleLeuSerMetLeuProLeuGluGlyLysCysLeuLeu 40

DB 149 AAGGAGGAGCGCCCCGAGAGATAGTCTTTTCTTCTTATATGAGGAAATCAGTCTG 208

Qy 41 GluLeuGlyAlaGlyIleGlyArgPheThrGlyGluLeuAlaGluLeuAlaGlyGlnVal 60  
Db 209 GAGCTTGGCGCGGAATAGTCTTACTGGAGAACTGGCTAAGACAGCTGGGCGATGTT 268  
Qy 61 IleAlaLeuAspPheIleGluSerAlaIleLeuLeuAlaGluValIleAsnGlyHisTyr 80  
Db 269 CTGCAATGGATTTTCATGGAGTGTGATTAAGAGATGAAGCATTAATGGCCATTAC 328  
Qy 81 LysAsnValLysPheMetCysAlaAspValThrSerProThrIleuSerPheProHis 100  
Db 329 GAAATGTCATCTTCATGTGTGCTGATGTATACATCTCCAGACCTGGTGATTTAGGACAA 388  
Qy 101 SerLeuAspValIlePheSerAsnTyrLeuLeuMetTyrLeuSerAspGluValGlu 120  
Db 389 TCGATTGATCTCATATTTTCAAACTGGCTACTGATGTATCTTTCAGACGAGGAGTCGAG 448  
Qy 121 AsnLeuValGluArgMetLeuValThrLeuLysProGlyGlyTyrIlePheArgGlu 140  
Db 449 AAGCTTGTAGAAAGAAATGTTAAATGGCTCAAGTTGGTGGCCCAATCTCTTAGAGAA 508  
Qy 141 SerCysPheHisGlnSerGlyAspHisLysArgLysSerAsnProThrHisTyrArgGlu 160  
Db 509 TCATGCTTCCATCATCTGGAGACTCAAAAAGGAAAGTGAATCCGACACATTATCGTGAA 568  
Qy 161 ProArgPheTyrThrLys 166  
Db 569 CCAGGTTTTACACTAG 586

Search completed: August 2, 2004, 00:48:18  
Job time : 552 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 1, 2004, 23:12:39 ; Search time 5580 Seconds  
(without alignments)  
3674.060 Million cell updates/sec

Title: US-10-031-331b-40

Perfect score: 2515  
Sequence: 1 HTVDLTIEAMLDQASDL.....KLRRKANGQRMGLFVAKKK 473

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 04  
Maximum Match 1004  
Listing first 45 summaries

Command line parameters: -DEV=xlh

-MODELS=frame+ p2n.model -DB=GenBml -QFMT=fastcap -SUFFIX=ice -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=500 -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODBS=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10031331 @CGN 1 1 3508 -runat 26072004 120351 14978 -NCFU=6 -ICPU=3  
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DBLEXT=7

Database :

GenBml:  
1: gb\_ba:  
2: gb\_hg:  
3: gb\_in:  
4: gb\_om:  
5: gb\_ov:  
6: gb\_pat:  
7: gb\_ph:  
8: gb\_pl:  
9: gb\_pt:  
10: gb\_ro:  
11: gb\_sts:  
12: gb\_sy:  
13: gb\_un:  
14: gb\_vi:  
15: em\_ba:  
16: em\_fun:  
17: em\_hum:  
18: em\_in:  
19: em\_mu:  
20: em\_om:  
21: em\_ov:  
22: em\_or:  
23: em\_pat:  
24: em\_ph:  
25: em\_pl:  
26: em\_ro:  
27: em\_sts:  
28: em\_un:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2515	100.0	1602	6	BD084055 Enviro
2	2515	100.0	1602	6	BD093370 Enviro
3	2515	100.0	1779	8	AB080186 Sueda ja
4	2204	87.6	2235	6	AX250551 Sequence
5	2204	87.6	2252	8	AF237633 Spinacia
6	2077	82.6	1459	8	AY133811 Arabidops
7	2077	82.6	1656	8	AF428454 Arabidops
8	2077	82.6	1660	8	AY063866 Arabidops
9	2063	82.0	1794	8	AF328858 Lycopersi
10	2044.5	81.3	1870	8	AB090883 Aster tri
11	2018	80.2	1834	8	AY136372 Arabidops
12	2011	80.0	1476	8	AY091683 Arabidops
13	2011	80.0	1770	8	AY058175 Arabidops
14	2011	80.0	1809	8	AF367299 Arabidops
15	2011	80.0	1878	8	AF197940 Arabidops
16	2009	79.9	1750	8	AY139479 Brassica
17	2004	79.7	1911	8	AY093093 Arabidops
18	1964	78.1	2069	8	AK069137 Oryza sat
19	1918.5	76.3	1782	8	AY065971 Trilicium
20	1680.5	66.8	29508	8	AC079676 Arabidops
21	1678	66.7	95295	8	AC020889 Genomic s
22	1561.5	62.1	74968	8	AB019230 Arabidops
23	1558	61.9	3221	8	AK102037 Oryza sat
24	1481	58.9	141715	8	AC130725 Oryza sat
25	1481	58.9	15911	2	AC136492 Oryza sat
26	1424	56.6	16686	8	AP003248 Oryza sat
27	1270	50.5	1120	6	AX250553 Sequence
28	1133.5	45.1	11754	2	AC138199 Medicago
29	1109	44.1	105937	8	AC012679 Arabidops
30	979	38.9	144305	2	AC137078 Arabidops
31	888.5	35.3	724	8	FVU234432 Hordeum v
32	405	16.1	286	6	AR252133 Sequence
33	389	15.5	1220	3	AY429590 Plasmodu
34	389	15.5	110000	2	PFMAL13.16 Continuation (17 o
35	363	14.4	280	6	AR251334 Sequence
36	347	13.8	274	6	AR248084 Sequence
37	330	13.1	260	6	AR245517 Sequence
38	329	13.1	31226	3	U39998 Caenorhabdi
39	296	11.8	349498	1	AP003002 Mesorhizo
40	286.5	11.5	23614	3	U64834 Caenorhabdi
41	288.5	11.5	140702	2	AC006888 Caenorhab
42	262	10.4	174	6	AR250330 Sequence
43	257.5	10.2	274	6	AR245829 Sequence
44	241.5	9.6	846	6	AX598613 Sequence
45	241.5	9.6	52101	6	AX598593 Sequence

# ALIGNMENTS

RESULT 1

BD084055	BD084055	1602 bp	DNA	linear	PAT 27-AUG-2002
LOCUS	BD084055	1602 bp	DNA	linear	PAT 27-AUG-2002
DEFINITION	Environmental stress-tolerant gene.				
ACCESSION	BD084055				
VERSION	BD084055.1	GI:22629665			
KEYWORDS	JP 2001333784-A/20.				
SOURCE	unidentified				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 1602)				
AUTHORS	Yamada, A., Ozeki, Y. and Saito, T.				
TITLE	Environmental stress-tolerant gene				
JOURNAL	Patent: JP 2001333784-A 20 04-DEC-2001;				
COMMENT	JAPAN SCIENCE AND TECHNOLOGY CORP Sueada japonica JP 2001333784-A/20 PD 04-DEC-2001 PI 19-JUL-2000 JP 2000219649 PF AKIYO YAMADA, YOSHIHIRO OZEKI, TAKEO SAITO PC C12N15/09, A01H5/00, C07K14/415, C07K16/16, C12N5/10, C12P21/02// PC C12N1/21, PC C12P21/08, C12N15/00, C12N5/00 CC Environmental stress-tolerant gene FH Key FT CDS Location/Qualifiers				
FEATURES	1..1602				
source	/organism="unidentified" /mol_type="genomic DNA" /db_xref="taxon:32644"				
ORIGIN					
Alignment Scores:	8,476-225	Length:	1602		
Pred. No.:	2515.00	Matches:	473		
Score:	100.00%	Conservative:	0		
Percent Similarity:	100.00%	Mismatches:	0		
Best Local Similarity:	100.00%	Indels:	0		
Query Match:	100.00%	Gaps:	0		
DB:	6				
US-10-031-331B-40 (1-473) x BD084055 (1-1602)					
QY	1	HisThrValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp	20		
Db	1	CACACGGTGTATTAACCACTTGAAGCTATGATGCTCGATCTCAAGCTCTCGATCTTAC	60		
QY	21	LysGluGluArgProGluIleLeuSerMetLeuProProLeuGluGlyLysCysCysLeuLeu	40		
Db	61	AAAGAGACAGCTCTGAGATTCCTTCAATGCTTCGCTTGAAGGAAATCCCTCTTG	120		
QY	41	GluLeuGlyAlaGlyIleGlyArgPheThrGlyGluLeuAlaGluLysAlaGlyGlnVal	60		
Db	121	GAACCTGGGGCTGGTATGCTCGTTTACTGGTGAATTCGCTGAGAAAGCTGGCCAGTT	180		
QY	61	IleAlaLeuAspPheIleGluSerAlaIleValLysLeuGlnValIleAsnGlyHisTyr	80		
Db	181	ATTGCTCTGATTCATGAGTGTCTATCAGAGATATGAGTATCATATGGGCACTAC	240		
QY	81	LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProHis	100		
Db	241	AAAAATGTCAGTTTATGCTGCTGATGTGACTTCTCCCACTCTCAGTTTCCCAACAT	300		
QY	101	SerLeuAspValIlePheSerAsnTrpLeuMetTyrLeuSerAspGluGluValGlu	120		
Db	301	TCATTGGATGTATCTCCATTTGTTACTCATGTATCTTCTCATGATGATGAGGTGAA	360		
QY	121	AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyGlyTyrIlePhePheAspGlu	140		
Db	361	AAATTGGTGAAGAAATTTGAATGTGAGCCAGGGGGTACATTTCTTCAGAGAA	420		
QY	141	SerCysPheHisGlnSerGlyAspHisLysArgLysSerAsnProThrHisTyrArgGlu	160		
Db	421	TCTTGTTCCTCATCTGGGATCAACACGAAAGCAATCCCAACCCCTACCTGCGTAA	480		
QY	161	ProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSer	180		
Db	481	CCTAGGTTCTACACTAAGCCCTTCAAGAGTGTCTATTTCAGAGTGGATCTGGAAATCT	540		
QY	181	TyrGluLeuSerLeuLeuSerCysLysCysIleGlyAlaTyrValArgAsnLysLysAsn	200		
Db	541	TATGAGCTCTCCCTACTAGTGCCTTGTGAGGCTTATGTGCAGAAACACAGAAAC	600		
QY	201	GlnAsnGlnIleSerTrpLeuTrpGlnLysValAspSerLysAspAspLysGlyPheGln	220		
Db	601	CAGAACCAAGATTAGTTGGTTGGTGGCAAAAGTTGATTTTCTAGGATGATAGGGGTTCC	660		
QY	221	ArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArgValPhe	240		
Db	661	CGATTTCTGGATCTAGCCAGTACAAGTGTATAGCATTTCTGCGATATGAGCGTGTATT	720		
QY	241	GlyProGlyTyrValSerThrGlyGlyTyrGluThrThrLysGluPheValSerMetLeu	260		
Db	721	GGCTCTGTTTATGTTAGCTCTGAGGATGATGAACCAACCAAGAGTGTGTGTCAATGCT	780		
QY	261	AspLeuLysProGlyGlnLysValLeuAspValGlyCysGlyIleGlyGlyAspPhe	280		
Db	781	GACTTGAAGCCCTGGCCAGAAAGGCTCTGAGTGTGGTGTGGATTTGGTGGAGGTGACTT	840		
QY	281	TyrMetAlaGluThrPheAspValGluValValGlyPheAspLeuSerValAsnMetIle	300		
Db	841	TACATGGCGGAGACCTTTGATGTTGGGTTTGGATTTGATCTCTCCGTTTAATATGATT	900		
QY	301	SerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluValAlaAsp	320		
Db	901	TCCTTTGCCCTTGAGCGTTCTATTGGCTTAATATGCTGCTGTTGAGTTGAGGTAGCAG	960		
QY	321	CysThrLysIleAsnTyrProAspAsnSerPheAspValIleTyrSerArgAspThrIle	340		
Db	961	TGCACCAAGATAAATACCTCTGATTAATCTTTTGTATGATCATCTATAGCCGTGACCAT	1020		
QY	341	LeuHisIleGlnAspLysProAlaLeuPheArgSerPheTyrIleTyrLeuLysProGly	360		
Db	1021	CTGCATATTACGACACAGCCCTGCTGTTTGTAGATCTCTTACAAATGTTGAGCCAGGA	1080		
QY	361	GlyLysValLeuIleSerAspTyrCysLysAlaGlyProProSerProGluPheAla	380		
Db	1081	GGTAAAGTTCTAATCAGTGATTTCTGCAGAAAGCTGGTCCACCTCACTGATTCGCC	1140		
QY	381	AlaTyrIleLysGlnArgGlyTyrAspLeuHisAspValLysGluTyrGlyGlnMetLeu	400		
Db	1141	GCTTACATTAGCAGAGGGGATGATCATCTCCATCATGATAAAGGAATATGGGAGATGCT	1200		
QY	401	LysAspAlaGlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIleArgVal	420		
Db	1201	AAAGATGCTGGATTTGTTGATGTTCTTCCGAGGATAGACTGAGCAGTTCATTCCAGTT	1260		
QY	421	LeuArgGlyGluLeuGluThrValGluLysAspValPheIleSerAspPheSer	440		
Db	1261	CTACGGAAGAACTAGAGCTGTTGAGAAAGGAAGGATGTGTTTCATTAGTATTTCTCT	1320		
QY	441	GluGluAspTyrAsnAspIleValGlyGlyTrpAsnAspLysLeuArgArgThrAlaLys	460		
Db	1321	GAGGAGCATTTACATTCACATTTGTTGGAGGTTGGAATGATTAAGTTGCGAGGACTGCC	1380		
QY	461	GlyGluGlnArgTrpGlyLeuPheValAlaLysValLys	473		
Db	1381	GGTGAGCAACGATGGGGCTGTGTTGTTGCCAAGAAAGAG	1419		
RESULT 2					
BD093370	BD093370	1602 bp	DNA	linear	PAT 27-AUG-2002
LOCUS	BD093370	1602 bp	DNA	linear	PAT 27-AUG-2002
DEFINITION	Environmental stress-tolerant gene.				
ACCESSION	BD093370				
VERSION	BD093370.1	GI:22638958			
KEYWORDS	WO 016006-A/20.				
SOURCE	unidentified				

ORGANISM unidentified  
unclassified.  
REFERENCE 1 (bases 1 to 1602)  
AUTHORS Yamada,A., Ozeki,Y. and Saito,T.  
TITLE Environmental stress-tolerant gene  
JOURNAL Patent: WO 0106006-A 20 25-JAN-2001;  
JAPAN SCIENCE AND TECHNOLOGY CORP, AKIYO YAMADA, YOSHIHIRO OZEKI,  
TAKEO SAITO  
COMMENT OS Suaeda japonica  
PN WO 0106006-A/20  
PD 25-JAN-2001  
PF 19-JUL-1999 JP 99P 235910, 24-MAR-2000 JP ODP 085377 PI  
AKIYO YAMADA, YOSHIHIRO OZEKI, TAKEO SAITO  
PC C12Q1/68, C12N15/29, C07K14/415, C07K16/16, C12P21/02, A01H5/00 CC  
FH Key Location/Qualifiers  
PT CDS (1). (1419).  
FEATURES  
source  
1..1602  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"  
ORIGIN  
Alignment Scores:  
Pred. No.: 8.47e-225 Length: 1602  
Score: 2515.00 Matches: 473  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-10-031-331B-40 (1-473) x BD093370 (1-1602)  
QY 1 HisThrValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp 20  
DB 1 CACACCGTGTGATTAACCAATGAGCTATGATGCTCGATTCATCAAGCTTCATGATCTGAC 60  
QY 21 LysGluGluArgProGluIleLeuSerMetLeuProProLeuGluGlyLysCysLeuLeu 40  
DB 61 AAGAGAGAACCTGCTGAGATCTCTTCAATGCTTCGCGCTCTTGAAGGAAAATGCTCTTG 120  
QY 41 GluLeuGlyAlaGlyIleGlyArgPheThrGlyGluLeuAlaGluLysAlaGlyGlnVal 60  
DB 121 GAACCTGGGGCTGGTATGCTGCTGCTTACTGGTGAATGGCTGAGAAAGCTGGCCAGGTT 180  
QY 61 IleAlaLeuAspPheIleGluSerAlaIleLysLysAsnGluValIleAsnGlyHisTyr 80  
DB 181 ATTGCTCTGGATTCATTGAGAGTGTCTATCAGAGAGATGAAGTAAATCAATGGGCACTAC 240  
QY 81 LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProHis 100  
DB 241 AAAAATGTCAAGTTATGT 120  
QY 101 SerLeuAspValIlePheSerAsnThrLeuMetTyrLeuSerAspGluGluValGlu 120  
DB 301 TCAITGGATGTGATATTCCTCAATGCTTACTCAATGATATCTTCTGATGAAGAGTGGAA 360  
QY 121 AsnLeuValGluArgMetLeuLysThrLeuLysProGlyGlyTyrIlePhePheArgGlu 140  
DB 361 AATTGGTGTGAAGAATGTGAAGAATGTGAAGAATGTGAAGAATGTGAAGAATGTGAAGAAT 420  
QY 141 SerCysPheHisGlnSerGlyAspHisLysArgLysSerAsnProThrHisTyrArgGlu 160  
DB 421 TCTGTGTTCCATCAATCTGGGGATCACAAAGCAAAAGCAATCCCAACCATACCGTGA 480  
QY 161 ProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSer 180  
DB 481 CTTAGGTTCTACATAAGGCTTCAAGAGTGTCAATTTGCAAGTGGATCTGGAACTCT 540  
QY 181 TyrGluLeuSerLeuLeuSerCysLysCysIleGlyAlaTyrValArgAsnLysAsn 200  
DB 541 TATGAGCTCTCTTACTAGTCAAAATGATTTGGAGCTTATGTCAGAAACAACAAAAC 600

QY 201 GlnAsnGlnIleSerTrrpLeuTrrpGlnLysValAspSerLysAspAspLysGlyPheGln 220  
DB 601 CAGAACCCAGATAGTGGTGGTGGCAAAAAGTTCATTTCTAAGCATGATAAGGGTTCAG 660  
QY 221 ArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArgValPhe 240  
DB 661 CGATTTCTGGATACTAGCCAGTACAAGTGTATAGCATCTCGGATATGAGGGTGTATT 720  
QY 241 GlyProGlyTyrValSerThrGlyGlyTyrGluThrThrLysGluPheValSerMetLeu 260  
DB 721 GGCCCTGGTATGTTAGCATCTGGAGGATAGAACCCCAAGAGTGTGTGTCAATGCTG 780  
QY 261 AspLeuLysProGlyGlnLysValLeuAspValGlyCysGlyIleGlyGlyLysAspPhe 280  
DB 781 GACTTGAAGCCCTGGCCAGAGGCTCTGGATGTGTGGTGTGTGGAAITGGTGGAGGTGACTT 840  
QY 281 TyrMetAlaGluThrPheAspValGluValGlyPheAspLeuSerValAsnMetIle 300  
DB 841 TACATGGCGGAGACCTTTGATGTTGAGTGTGTGGATTTGATCTCTCGTTAATATGATT 900  
QY 301 SerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluValAlaAsp 320  
DB 901 TCCTTTGGCCCTTGAGCGCTTCTATTGGGCTTAAATGGCTGTGTGAGTGTGAGGTAGCAGAT 960  
QY 321 CysThrLysIleAsnTyrProAspAsnSerPheAspValIleTyrSerArgAspThrIle 340  
DB 961 TGCACCAAGATAAACTACCTGATAACTCTTTTGTGATCTATAGCCGTGACACCAATT 1020  
QY 341 LeuHisIleGlnAspLysProAlaLeuPheArgSerPheTyrLysTrrpLeuLysProGly 360  
DB 1021 CTGCATATTCAAGACAGGCTGGGTTGTTAGATCTTCTAACAATGTTGAGGCCAGGA 1080  
QY 361 GlyLysValLeuIleSerAspTyrCysLysLysAlaGlyProProSerProGluPheAla 380  
DB 1081 GGTAAAGTGTCTAAATAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
QY 381 AlaTyrIleLysGlnArgGlyTyrAspLeuHisAspValLysGluTyrGlyGlnMetLeu 400  
DB 1141 GCTTACATTAAAGCAGAGGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200  
QY 401 LysAspAlaGlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIleArgVal 420  
DB 1201 AAGATGCTGGATTTGTGATGTTCTTGGCCGAGGATAGAACTGAGCAGATTCATTCGAGTT 1260  
QY 421 LeuArgLysGluLeuGluThrValGluLysGluLysAspValPheIleSerAspPheSer 440  
DB 1261 CTACGGAGGAACTAGAGACTGTTCAGAGAGAAAGGATGTTGTTTCAATTAGTATTCTCT 1320  
QY 441 GluGluAspTyrAsnAspIleValGlyGlyTrrpAsnAspLysLeuArgGlnThrAlaLys 460  
DB 1321 GAGGAGGATTCATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380  
QY 461 GlyGluGlnArgTrrpGlyLeuPheValAlaLysLysLys 473  
DB 1381 GGTGAGCAGCATGGGGTCTGTTCGTTGCCAAGAGAGAG 1419  
RESULT 3  
AB080186  
LOCUS Suaeda japonica PEAMT mRNA linear PLN 20-FEB-2003  
DEFINITION N-methyltransferase, complete cds.  
ACCESSION AB080186  
VERSION AB080186.1 GI:28436073  
KEYWORDS  
SOURCE Suaeda japonica  
ORGANISM Suaeda japonica  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Caryophyllales; Amaranthaceae; Suaeda.  
REFERENCE 1  
AUTHORS Yamada,A., Nozawa,G.T., Tanimoto,S. and Ozeki,Y.  
TITLE Glycinebetaine synthesis in Suaeda japonica

Unpublished  
2 (bases 1 to 1779)  
Yamada, A., Nozawa, G.T. and Ozeki, Y.  
Direct Submission  
Submitted (19-FEB-2002) Akiyo Yamada, Tokyo University of  
Agriculture and Technology, Department of Biotechnology, Naka-cho  
2-24-16, Koganei, Tokyo 184-8588, Japan  
(E-mail: yamadenec@at.tuat.ac.jp, Tel: 81-423-88-7239,  
Fax: 81-423-88-7239)  
Location/Qualifiers  
1. 1779  
/organism="Suaeda japonica"  
/mol\_type="mRNA"  
/db\_xref="taxon:90346"  
/country="Japan: Saga, Ariake Sea"  
1. 1779  
/gene="PEANT"  
115. 1599  
/gene="PEANT"  
/codon\_start=1  
/product="phosphoethanolamine N-methyltransferase"  
/protein\_id="BAC57432.1"  
/db\_xref="GI:28436074"  
/translation="MAASGMEIQRBRDVPKVVHETVDTLATEAMLDQASDLKKEE  
RPEILSMLEKLCLELGGAGRTGRLAEGAGVIALQFIRSAIKNEVINCHYK  
NVMPCADVTSPILSPHSLDVI FSNMILMYLSDEEVNLEWMLKPGVYIPR  
ESCFQSGDHKRSNPTHTYRPFYTNKAECHLQDGSNSVELSLCKICIGAVN  
KKNQIQISMLQKVDSDKDGQFPLDTQYKNSILRYRFRVFGYVSTGYETIKE  
FVSLDLKQKQVLDVCGGIGGDFYNAETFDVWVGFSLVNMILKPGGKVLISGLKCA  
VEFVADCTKINPTDVSFDVIYSRDTLHIQKPALEFRSPYKMLKPGGKVLISGLKCA  
AGPSPFAAYIKQGYDLHVKYGMKLDAGFDVLAEDTEQFIRVLRKELETVE  
KEKDVFIQDFSEEDYNDIVGMDKLRRRTANGKQWGLFVAKK"

ORIGIN

Alignment Scores:  
Pred. No.: 9,726-225 Length: 1779  
Score: 2515.00 Matches: 473  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0

US-10-031-331b-40 (1-473) x AB080186 (1-1779)

Qy 1 HistThraspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp 20  
Db 178 CACACCGTGTGATTAACCTTTGAAGCTATGATGCTGCTTCTCAAGCTTCTGATCTTGC 237  
Qy 21 LysGluGluArgProGluLeuLeuSerMetLeuProLeuGluGluLysCysLeuLeu 40  
Db 238 AAGAAGAACGCTCTGAGATCTTTCAATGCTTCCGCTCTTGAAGGAAATGCTCTTTC 297  
Qy 41 GluLeuGluAlaGluIleGluArgPheThrGluGluLeuAlaGluLysAlaGluGlnVal 60  
Db 298 GAACCTTGGGCTGGTATGCTGCTTCTTCAATGCTTCCGCTCTTGAAGGAAATGCTCTTTC 357  
Qy 61 IleAlaLeuAspPheIleGluSerAlaIleLysLysAsnGluValIleLeuGlyIleVal 80  
Db 358 ATTGCTCTGATTTCAATGAGAGTCTATCAAGAGATGAGATGAGTATCAATGAGGACTAC 417  
Qy 81 LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProHis 100  
Db 418 AAAAAATGTCAGTTTATGCTGATGCTGATGCTCTCCACTCTCAGTTTCCACCCACAT 477  
Qy 101 SerLeuAspValIlePheSerAsnTrpLeuMetTyrLeuSerAspGluGluValGlu 120  
Db 478 TCATTGATGATGATATTCATATGCTGATGCTGATGCTGATGCTGATGATGATGATGATG 537  
Qy 121 AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyGlyTyrIlePhePheArgGlu 140  
Db 538 AATTGCTTGAAGAGATGTTGAAGATGTTGAAGATGTTGAAGATGTTGAAGATGTTGAAGAT 597  
Qy 141 SerCysPheHisGlnSerGlyAspHisLysArgLysSerAsnProThrHisLysArgGlu 160

Db 598 TCTGTGTTCCATCAATCTGGGATCACAAGCAAAAGCAATCCACCACCTACCGTGA 657  
Qy 161 ProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSer 180  
Db 658 CCTAGGTTCTTACACTAAGGCTTCAAGAGATGCTATTCTCAAGATGGATCTGGAACTCT 717  
Qy 181 TyrGluLeuSerLeuLeuSerCysLysCysGlyIleGlyAlaTyrValArgAsnLysLys 200  
Db 718 TATGAGCTCTCCCTACTTACCTGCAATGATATGAGGCTTATGTCAGAAACAGAAAC 777  
Qy 201 GlnAsnGlnIleSerTrpLeuTrpGlnLysValAspSerLysAspAspLysGlyPheGln 220  
Db 778 CAGAACCAAGATTAGTTGGTGTGGCAAAAAGTTGATTTCTAAGGATGATGAAGGGTTCAG 837  
Qy 221 ArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArgValPhe 240  
Db 838 CGATTTCTGGATACTAGCCAGTACAGTGTATAGCAATCTGGGATATGAGGGTGTATTT 897  
Qy 241 GlyProGlyTyrValSerThrGlyGlyTyrGluThrThrLysGluPheValSerMetLeu 260  
Db 898 GGCCCTGGTATGTTAGCACTGGAGGATATGAACCAACCAAGAGTTTGTGTCAATGCTG 957  
Qy 261 AspLeuLysProGlyGlnLysValLeuAspValGlyCysGlyIleGlyGlyAspPhe 280  
Db 958 GACTTGAAGCTCTGCCAGAGAGTCTCTGGATGTTGGTGTGGAAATTTGGTGGAGGTGACTTT 1017  
Qy 281 TyrMetAlaGluThrPheAspValGluValValGlyPheAspLeuSerValAsnMetIle 300  
Db 1018 TACATGGCGAGACCTTTGATGTTGAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 1077  
Qy 301 SerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluValAlaAsp 320  
Db 1078 TCCTTTGGCTTGGAGCGTCTTATTTGGCTTAAATGCTGTGTTGAGTTGAGGATGACAT 1137  
Qy 321 CysThrLysIleAsnTyrProAspAsnSerPheAspValIleTyrSerArgAspThrIle 340  
Db 1138 TGCACCAAGATAAATACCTACCTGATTAATCTTTTGTATGTCATCTATAGCCGTGACACCAT 1197  
Qy 341 LeuHisIleGlnAspLysProAlaLeuPheAspSerPheTyrIleTrpLeuLysProGly 360  
Db 1198 CTGCATATTCAGGACAGCCCTGCTGTTTGTGATCTTCTACAAATGTTGAGGACGGA 1257  
Qy 361 GlyLysValLeuIleSerAspTyrCysLysAlaGlyProProSerProGluPheAla 380  
Db 1258 GGTAAGTTCTTAATCAGTGATTAATGCAAGAAAGCTGGTCCACCTCACCCTGATTCGCT 1317  
Qy 381 AlaTyrIleLysGlnArgGlyTyrAspLeuHisAspValLysGluTyrGlyGlnMetLeu 400  
Db 1318 GCTTACATTAAGCAGAGGGGATGATGATCTCAATGATTAAGGATATGAGGATGCTT 1377  
Qy 401 LysAspAlaGlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIleArgVal 420  
Db 1378 AAGATGCTGATTTGTTGATGTTCTTGGAGGATGAGAACTGAGCAGTTCATTCAGTT 1437  
Qy 421 LeuArgLysGluLeuGluThrValGluLysGluLysAspValPheIleSerAspPheSer 440  
Db 1438 CTACGAGGAGGACTAGACACTGTTGAGAGGAAAGGATGTTGTTTATGATGATTTCTCT 1497  
Qy 441 GluGluAspTyrAspAspIleValGlyTyrAsnAspLysLeuArgThrAlaLys 460  
Db 1498 GAGGAGGATTAACATGACATTTGTTGAGGATTTGGATGATGATGATGATGATGATGAT 1557  
Qy 461 GlyGluGlnArgTyrGlyLeuPheValAlaLysLysLys 473  
Db 1558 GGTGAGCAACGATGGGGTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1596

RESULT 4  
AX250551 LOCUS 2235 bp DNA linear PAT 05-OCT-2001  
AX250551 Sequence  
AX250551 DEFINITION  
AX250551.1 GI:15984289  
AX250551.1 GI:15984289

## KEYWORDS

SOURCE Spinacia oleracea (spinach)  
ORGANISM Spinacia oleracea

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
TITLE Caryophyllales; Amaranthaceae; Spinacia.

Hanson, A.D., Nuccio, M.L., and Henry, S.A.  
S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase  
compositions and methods for modulating lipid biosynthesis in  
plants

JOURNAL Patent: WO 0168870-A 1 20-SEP-2001;

FEATURES University of Florida (US); Carnegie-Mellon University (US)

## source

Location/Qualifiers  
1..2235  
/organism="Spinacia oleracea"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3562"

## ORIGIN

Alignment Scores:  
Pred. No.: 1,31e-195 Length: 2235  
Score: 2204.00 Matches: 408  
Percent Similarity: 93.64% Conservatives: 34  
Best Local Similarity: 86.44% Mismatches: 30  
Query Match: 87.63% Indels: 0  
DB: 6 Gaps: 0

US-10-031-331b-40 (1-473) x AX250551 (1-2235)

QY 1 HisThrValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuApp 20  
DB 317 CACTCTGTGTGATTTGACTGTGTGGAGCTATGATCTTGAATTCACAGCTTCAGATCTTCAC 376  
QY 21 LysGluGluArgProGluIleLeuSerMetLeuProLeuGluGluGlyCysLeuLeu 40  
DB 377 AAAGTGGAGCGACCTGAGGTACTTTCATGCTTCCACCTTATGAGGAAAGTGTGTCTTA 436  
QY 41 GluLeuGlyAlaGlyIleGlyArgPheThrGlyGluLeuAlaGluGlyAlaGlyVal 60  
DB 437 GNACTCGGTGCTGTGTATGTGTCTTTACTGGTGAATTTGGCGAGAAAGCTAGCCAGTGC 496  
QY 61 IleAlaLeuAspPheIleGluSerAlaIleLysLeuAsnGluValIleAsnGlyHisTyr 80  
DB 497 ATCGCTCTGGATTTTCATTTGAGGTGTATTAAGAGAGATGAAGCATTAATGGGCATATC 556  
QY 81 LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProHis 100  
DB 557 AAAAATGTGAAGTTTATGTCTGTGATGTGACATCTCCAAAGTCTCAATTTCCCAAT 616  
QY 101 SerLeuAspValIlePheSerAsnTrpLeuLeuMetTyrLeuSerAspGluGluValGlu 120  
DB 617 TCGGTGATATCATTTCTCCAAATGGCTACTCATGTATCTTCTGATGAAGAGGTTGAG 676  
QY 121 AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyGlyTyrIlePhePheArgGlu 140  
DB 677 CGTCTGGTTCAAAGGATTTGAATGTGTGAAGCCAGGAGGATACATTTTCTTCAGAGAA 736  
QY 141 SerCysPheHisGlnSerGlyAspHisLysArgLysSerAsnProThrHisTyrArgGlu 160  
DB 737 TCTTGTGTTTTCATCAATCAGAGATATCAAGCGCAAAAGCAATCCAAACCATCCAGTGA 796  
QY 161 ProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSer 180  
DB 797 CCTAGGTTCTACACCAAGATCTTCAAGATGCCATATGCAAGATGATCTTGGGAGTCC 856  
QY 181 TyrGluLeuSerLeuSerCysGlyCysIleGlyAlaTyrValArgAsnLysLysAsn 200  
DB 857 TATGAGTCTCCCTTAATGTGCTGCAATGATTTGAGCTATGTCAAAAGCAGAGAGAT 916  
QY 201 GlnAsnGlnIleSerTrpLeuTrpGlnLysValAspSerLysAspAspLysGlyPheGln 220  
DB 917 CAGAACAGATAGCTGTGTTATGGCAGAAAGTGTGATTCAGAGGATGACAGGGGTTCCAG 976

QY 221 ArgPheLeuAspThrSerGlnTyrIleCysAsnSerIleLeuArgTyrGluArgValPhe 240  
DB 977 CGATTTCTTGATTTCTAGTCAATACAGATTTAAACAGCATACTGCGTTATGAGCGTGTATTT 1036  
QY 241 GlyProGlyTyrValSerThrGlyGlyTyrGluThrThrLysGluPheValSerMetLeu 260  
DB 1037 GGTCTGCTTATGTAGTACCGAGGAGCTCCAAACCAACCAAGAGTTTGTATCAAGCTT 1096  
QY 261 AspLeuLysProGlyGlnLysValLeuAspValGlyCysGlyIleGlyGlyAspPhe 280  
DB 1097 GACTTGAAGCTTGGCCAGAGGTCTTAAGTGTGGTGTGGCATAGTGGAGTGATTTT 1156  
QY 281 TyrMetAlaGluThrPheAspValGluValValGlyPheAspLeuSerValAsnMetIle 300  
DB 1157 TACATGGCAGAGAACTATGATGTTGAGGTTGTTGGAATGATCTCTCCATTAATATGAT 1216  
QY 301 SerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluValAlaAsp 320  
DB 1217 TCTTTTGGCTTGGAGCGCTCAATTTGGCTCAATATGCTGTGAGTTTGGAGTGGAGAT 1276  
QY 321 CysThrLysIleAsnTyrProAspAsnSerPheAspValIleTyrSerArgAspThrIle 340  
DB 1277 TGACCAAGAAAGATTAACCTGAAACTCTTTTGAATGTCATCTACAGCGGTGATACCAT 1336  
QY 341 LeuHisIleGlnAspLysProAlaLeuPheArgSerPheTyrLysTyrIleLysProGly 360  
DB 1337 CTGCATATTCAGGACAAACCTGCTTTATTTAGATCTTCCACAAATGTTGAACTTGA 1396  
QY 361 GlyLysValLeuIleSerAspTyrCysValLysAlaGlyProProSerProGluPheAla 380  
DB 1397 GCGAAGTCTTATTTAGTACTTGTAGAGTGTGTGACACCTTCAGCTGAATTTGCT 1456  
QY 381 AlaTyrIleLysGlnArgGlyTyrAspLeuHisAspValLysGlyTyrGlyGlnMetLeu 400  
DB 1457 GCATACATTCAGGACAGGAGGATATGATCTCCAGATGTGAAGCATATGGCAAGATGCT 1516  
QY 401 LysAspAlaGlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIleArgVal 420  
DB 1517 AAGATGCTGGATTCGTGTAGGTTATTTGTGAGATAGGATGACCTTCATTCATGATTTCT 1576  
QY 421 LeuArgLysGluLeuGluThrValGluLysGluLysAspValPheIleSerAspPheSer 440  
DB 1577 CTGCAGAGGAGTATGATGCTCTTGAACAGGAGAGAGATGATCTTCATTCATGATTTCTCT 1636  
QY 441 GluGluAspTyrAsnAspIleValGlyTyrAsnAspLysLeuArgArgThrAlaLys 460  
DB 1637 GAGGAGGATTTATAACGACATAGTTGATGTTGAGGAGGAGGAGTGTGCTGAGGACTACAG 1696  
QY 461 GlyGluGlnArgTyrGlyLeuPheValAlaLysLys 472  
DB 1697 GGTGAGCAACATGGGTTTGTTCATTTGCCAAGAAA 1732

## RESULT 5

AP237633

LOCUS

DEFINITION

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Spinacia oleracea (spinach)

Spinacia oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Caryophyllales; Amaranthaceae; Spinacia.

1 (bases 1 to 2252)

Nuccio, M.L., Ziemak, M.J., Henry, S.A., Weretilnyk, E.A. and

Hanson, A.D.

CDNA cloning of phosphoethanolamine N-methyltransferase from

spinach by complementation in *Schizosaccharomyces pombe* and

characterization of the recombinant enzyme

J. Biol. Chem. 275 (19), 14095-14101 (2000)

AP237633 2252 bp mRNA linear PLN 10-MAY-2000  
Spinacia oleracea phosphoethanolamine N-methyltransferase (PEMT)  
mRNA, complete cds.

AP237633

AP237633.1 GI:7407188



MEDLINE 20261526  
EUNRD 10799484  
2 (bases 1 to 2252)  
Ruccio, M.L. and Hanson, A.D.  
Direct Submission  
Submitted (21-FEB-2000) Horticultural Sciences, University of  
Florida, P.O. Box 110690, Gainesville, FL 32611, USA  
JOURNAL Location/Qualifiers  
FEATURES  
1..2252  
/organism="Spinacia oleracea"  
/mol\_type="mRNA"  
/db\_xref="taxon:3562"  
/tissue\_type="sealined leaf"  
1..2252  
/gene="PEMT"  
254..1738  
/gene="PEMT"  
/EC\_number="2.1.1.103"  
/function="catalyzes S-adenosylmethionine dependent  
N-methylation of phosphoethanolamine,  
phosphomonomethyl ethanolamine and  
phosphodimethyl ethanolamine"  
/codon\_start=1  
/product="phosphoethanolamine N-methyltransferase"  
/protein\_id="AA061950.1"  
/db\_xref="GI:7407189"  
/translation="MAASMGVLOEVEVFKYWIHSVDLTVEAMMLDSQSLDKVE  
RPEVLSKLPYEGSKVLELGAGIGRTGELAKASQVIALQPIESVKNESLNGHYK  
NVKPMCAVDTPSINSPNSVDIIFSNMALLATLSDEEVEVERLVEKMLKPGGVIPR  
ESCFQSGDHRKSNPHTYEPFPIKIPFCHQMDDSGNSTLSLGCCKIGAYKS  
KKNQKQISLMQKVDSDDKGDFLDSSQYKFNISLRYERVFGPIVSTGGLETTKE  
FVSKLDLPQKQVLDVCGGGDFMAENYDVRVVGIDLSINMISPALEKSLGKCA  
VEFEVADCTKKDIPENSFOVYSRDTILHIDQKALPFRSFKMLKPGKVLISDYCK  
AGTSEAFAYIRQGVLDHVKAYGKMLKADGFEVIAENRTDQFIQVLQKELDLE  
QEKKDPIDDSSEVDYNDIVDMKAKLVRTTEGQKGLFLAKM"

ORIGIN  
Aligned Scores: 1.33e-195 Length: 2252  
Pred. No.: 2204.00 Matches: 408  
Score: 93.64% Conservative: 34  
Percent Similarity: 93.64%  
Best Local Similarity: 86.44% Mismatches: 30  
Query Match: 87.63% Indels: 0  
DB: 8 Gaps: 0

US-10-031-331b-40 (1-473) x AF237633 (1-2252)

Qy 1 HistHValAopLeuThrileGluAlaMetMetLeuAaspSerGlnAlaSerAenLeuAap 20  
Db 317 CACTCTGTGATTTGACCTGTTGAGGCTATGATGCTTCAACAGCTTCAGATTTGAC 376  
Qy 21 LysGluAuaGpProGluileLeuSerMetLeuProLeuGluGlyLysCysLeuLeu 40  
Db 377 AAGTGGAGCGACCTGAGTACTTCCATGCTTCCAGCTTATGAAGGAAGTCTGTCTTA 436  
Qy 41 GluLeuGlyAlaGlyileGlyAopPheThrileGluLeuAlaGluLysAlaGlyGlnVal 60  
Db 437 GAACCTCGGTCGTGATTTGCTGCTTACTGCTGATTTGCTGATTTGCTGATTTGCTG 496  
Qy 61 IleAlaLeuAapPheileGlySerAlaLysLysAlaLysLysLysLysLysLysLys 80  
Db 497 ATCGCTCTGGATTTCAATGAGAGTGTATTAAGGAAGATTAAGGAAGATTAAGGAAG 556  
Qy 81 LysAanValLysPheMetCysAlaAapValThrSerProThrLeuSerPheProPhe 100  
Db 557 AANAATGAGATTTATGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 616  
Qy 101 SerLeuAapValilePheSerAenTrpLeuLeuMetTrpLeuSerAapGluGluValGlu 120  
Db 617 TCGGTGATATCATATTCCTCAATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 676  
Qy 121 AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyLysTrpLeuPheAArgGlu 140

Db 677 CGTCTGGTGAAGAGATGTTGAATGGTTGAAGCCAGGAGATACATTTCTTTCAGAGAA 736  
Qy 141 SerCysPheHieGlnSerGlyAaspHisLysArgLysSerAenProThrHisLysArgGlu 160  
Db 737 TCTTGTGTTTTCATCAATCAGGAGATCAAGAGCGCAAGCAATCAACCCACTACCGTGA 796  
Qy 161 ProAArgPheTrpThrLysAlaPheLysGluCysHisLeuGlnAaspGlySerGlyAanSer 180  
Db 797 CTTAGGTTCTACACCAAGATCTTCAAGATGCCATATGCAAGATGATTTCTGGGAATCC 856  
Qy 181 TrpGluLeuSerLeuLeuSerCysLysCysLysGlyAlaValValArgAanLysLysAan 200  
Db 857 TATGAGCTCTCCCTAATTCCTGCAATGATTTGAGGATTTGATGCAAGCAAGCAAGAT 916  
Qy 201 GlnAanGlnileSerTrpLeuTrpGlnLysValAaspSerLysAaspLysGlyPheGln 220  
Db 917 CAGAACCAAGATGCTGTTATGGCAGAAAGTTGATTCAGAGGATGACAGGGGTTCCAG 976  
Qy 221 ArgPheLeuAaspThrSerGlnTrpLysCysAanSerLysLeuArgTrpGluArgValPhe 240  
Db 977 CGATTCCTGGATTTCTAGTCAATACAGTTTACACAGATCTGCGTTATGAGCGGTATTT 1036  
Qy 241 GlyProGlyTrpValSerThrGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 260  
Db 1037 GGTCTCTGTTATGTTAGTACCGGAGGCTCGCAACACCAAGAGTTTGTATCAAGCTT 1096  
Qy 261 AspLeuLysProGlyGlnLysValLeuAaspValGlyCysGlyLysGlyGlyGlyGly 280  
Db 1097 GACTTGAAGCTGCGCCAGAGAGTCTAGATGTTGGGTTGTGGCATAGGTGGAGGTGATTT 1156  
Qy 281 TyrMetAlaGluThrPheAapValGluValValGlyPheAaspLysValAanMetile 300  
Db 1157 TACATGGCAGGAAGTATGATGTTGAGTTGTTGGATTTGATCTCCATTAAATGATTT 1216  
Qy 301 SerPheAlaLeuGluAArgSerLysGlyLysLysCysAlaValGluPheGluValAa 320  
Db 1217 TCTTTGCGCTTGGAGCGCTCAATTTGGCTCAATGCTGTTGAGTTTGGAGTGGAGAT 1276  
Qy 321 CysThrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 340  
Db 1277 TGCACCAAGAGATTTACCTGTAAGAGTCTGTTGATGCTGATGCTGATGCTGATGCT 1336  
Qy 341 LeuHisLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 360  
Db 1337 CTGCATATTCAGGACCAACCTGCTTATTTAGATCTCTCACAAATGTTGAACTGGA 1396  
Qy 361 GlyLysValLeuLysSerAaspTrpCysLysLysLysLysLysLysLysLysLysLys 380  
Db 1397 GCGCAAGATTTCTATTATGAGTCTGTTAGAGTCTGTTAGAGTCTGTTAGAGTCTGTT 1456  
Qy 381 AlaTrpLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 400  
Db 1457 GCATACATTCAGGACAGGAGGATATGATCTCCAGTGTGAAGGCATATGCAAGATGCTT 1516  
Qy 401 LysAapAlaGlyPheValAaspValLeuAlaGluAaspArgThrGlnPheLysVal 420  
Db 1517 AAGATGCTGGAATTCGTTGAGGTTATGCTGAGATAGGACTGACCATGTTCTTCAAGTT 1576  
Qy 421 LeuArgLysGluLeuGluThrValGlyLysLysLysLysLysLysLysLysLysLys 440  
Db 1577 CTGCAAGGAAGTACTAGTCTCTTGCACAGGAGGATGATGATTCATTCATGATTTCTCT 1636  
Qy 441 GluGluAaspTrpAanAaspLysLysLysLysLysLysLysLysLysLysLysLysLys 460  
Db 1637 GAGGAGGATTTAAGCAGATAGTGTGATGTTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1696  
Qy 461 GlyGluGlnArgTrpGlyLeuPheValAlaLysLys 472  
Db 1697 GGTGAGCAACATGGGGTTGTTGTTCAATGCCAAGAA 1732  
RESULT 6  
AY133811  
LOCUS AY133811 1459 bp mRNA linear PLN 18-SEP-2002

**DEFINITION** Arabidopsis thaliana clone U1199 putative phosphoethanolamine N-methyltransferase (At1g48600) mRNA, complete cds.

**ACCESSION** AY133811

**VERSION** AY133811.1 GI:22136801

**KEYWORDS** FLI CDNA.

**SOURCE** Arabidopsis thaliana (thale cress)

**ORGANISM** Arabidopsis thaliana

**REFERENCE** Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; 1 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids; 1 (bases 1 to 1459)

**AUTHORS** Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Deng,J.M., Heuan,V.W., Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C., Wu,H.C., Yu,G., Yuan,S., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinzaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.

**TITLE** Arabidopsis Open Reading Frame (ORF) Clones

**JOURNAL** Unpublished

**REFERENCE** 2 (bases 1 to 1459)

**AUTHORS** Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Deng,J.M., Heuan,V.W., Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C., Wu,H.C., Yu,G., Yuan,S., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinzaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.

**TITLE** Direct Submission

**JOURNAL** Submitted (17-JUL-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

**COMMENT** The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinzaki,K.

The Salk, Stanford, PDEC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAPL cDNAs: Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Deng,J.M., Heuan,V.W., Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C., Wu,H.C., Yu,G., Yuan,S., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinzaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.

Yamada,K. (SSP/PDEC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinzaki,K. (RIKEN GSC) and Theologis,A. (SSP/PDEC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

**FEATURES** source

1..1459

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/db\_xref="taxon:3702"

/chromosome="1"

/clones="U1199"

/ecotype="Columbia"

/note="This clone is in pUNI 51."

1..1459

/gene="At1g48600"

1..1428

/gene="At1g48600"

/codon\_start=1

/evidence="experimental"

/product="putative phosphoethanolamine N-methyltransferase"

/protein\_id="AAW91745.1"

/db\_xref="GI:22136802"

/translations="MEHSSDLTVEAMMLDSKASLDKBERPEVLSLIPPVEGKSVLEL

CAGTGRGTGELAQAGAEVIALDFESAIQKESVNGHYKNIKPMCADVTSPLKIKOG  
SIDLIFSNLLMYLSDKEVLEMAERIMGMVKGPIYFFRESCHQSGSKSKSNPTHY  
REPRFYTKVFOCQTRDASNGSFLSMVCKICIGAYVKNQKQNOICMIWOKVSVEND  
KDFPFLDNVOYKSSGILRYERFVGGYVSTGGFTTKEFVAKMDLKPQOKVLDVGGC  
TGGSDPYMAENFDVHVGLDLSNMISFALERAIGLKSVPVADACTTKTVPNSSD  
VIYSRDTLHIHQKPAIFRTFFKWLKPGCKVLITDYCRSAETPSPEFARYLKORGYD  
HDQAYGQMLKADGPFDDVIAEDRTDQVQVLRLELKEVEKEEFISDFSEEDYNDIV  
GWSAKLERTASGQKGLFIADKK"

misc\_difference 282 /gene="At1g48600"

/note="compared to genomic sequence"

/replace="c"

1429..1459

/gene="At1g48600"

3'UTR

ORIGIN

Alignment Scores:

Pred. No.: 5.22e-184 Length: 1459

Score: 2077.00 Matches: 379

Percent Similarity: 90.06% Conservative: 47

Best Local Similarity: 80.13% Mismatches: 47

Query Match: 82.58% Indels: 0

DB: 8 Gaps: 0

US-10-031-331b-40 (1-473) x AY133811 (1-1459)

QY 1 HistHrValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp 20

DB 7 CAITCTAGTATTGACTGTTGAAGCTATGATGCTTGAAGCTCTTAAGCTCTTGAATTTGAC 66

QY 21 LysGluGluArgProGluIleLeuSerMetLeuProLeuGluGluGlyLysCysLeuLeu 40

DB 67 AAGAGAGACGCTCTGAGGTACTCTCTTTATCCCAACCATATGAGAGGGGAAATCTGTGCTT 126

QY 41 GluLeuGlyAlaGlyIleGlyArgPheThrGlyGluLeuAlaGluLysAlaGlyGlnVal 60

DB 127 GAACCTGGAGCTGCTATTGCTGCTTTTCACTGGTGAATTCGCTCAAAAGCTGGTGAAGTT 186

QY 61 IleAlaLeuAspPheIleGluSerAlaIleLysLysAsnGluValIleAsnGlyHisTyr 80

DB 187 ATCGCTCTGACTTCATCGAAGCGGATTCAGAGATGAAGTGTATTAATGGGCAATAC 246

QY 81 LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProHis 100

DB 247 AAGAACATCAAGTTATGTGTGCTGATGATATCCAGACTCCAGACTTGAAGATCAAGATGGA 306

QY 101 SerLeuAspValIlePheSerAsnTrpLeuLeuMetTyrLeuSerAspGluGluValGlu 120

DB 307 TCTATCGACTTGATTTTCTCAAACTGTTGCTCATGTATCTCTCTGATTAAGAGGTGAA 366

QY 121 AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyGlyTyrIlePhePheArgGlu 140

DB 367 CTATGGCAGAGAGAAATGATGATGGTCAAGCAGGGGGATACATTTCTTTCAGAGAA 426

QY 141 SerCysPheHisGlnSerGlyAspHisLysArgLysSerAsnProThrHisTyrArgGlu 160

DB 427 TCTTGTCTTCATCATCTGGGACAGCAGACGAAAGTCAAACTCATCTCATCCGTGAA 486

QY 161 ProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSer 180

DB 487 CCCAGATTCTACACAAAGGTTTCCAGGAAATGTCAGACACGCTGATGCTTCGGCATICA 546

QY 181 TyrGluLeuSerLeuSerCysLysCysIleGlyAlaTyrValArgAsnLysLysAsn 200

DB 547 TTGAGCTCTCTATGTTGGCTGCAAAATGCAATTTGGGCTTATGTGTAAGAACAGAGAAAT 606

QY 201 GlnAsnGlnIleSerTrpLeuTrpGlnLysValLeuAspSerLysAspLysGlyPheGln 220

DB 607 CAGAACTCAGATTGCTGGATATGGCAAAAGTCCAGCTGGAGAGATGACAGAGGATTTCAG 666

QY 221 ArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArgValPhe 240

DB 667 CGTTTCTTGACAAATGTTCAATACAAGTCTAGTGGGATCTTTGGCTATGACGCTGCTTT 726

241	QY	GlyProGlyLysValSerThrGlyCylTyrgluThrThrygluPheValserMetLeu	260
727	DB	GGGGAAGGATATGTUGACACTGGTGGATTTCAGACAACTAAGAGATTTTGCGCGAAGATG	786
261	QY	AspLeuLysProGlyGlnLysValLeuAspValGlyCysGlyIleGlyGlyLysAspPhe	280
787	DB	GACCTTAACCCGGGACAGAAAGTCTTAGATGTTGGTGTGTGATCGTGCGAGTGCACTTC	846
281	QY	TyrMetalaGluThrPheaspvalgluValValglyPheaspLeuSerValaemMetile	300
847	DB	TACATGGCTGAGAANTTCGATGTTCTCATGTGTGTGGATCCGATCTGTGGGCACATGATC	906
301	QY	SerPhealaleuGluArgSerIleGlyLeuLysCysAlavalGluPheGluValalashp	320
907	DB	TCATTTCGCACTGAGCGGGGCCATTGGCAATTCATAAATGCTCAGTCGAGTTTGAAGTCGCTGAT	966
321	QY	CysThrLysIleAsnTyrProaphenSerPheaspValilleTyrSerArgaspThrile	340
967	DB	TGCACCACCAAACATATCCCGATTAATTCCTTTGATGTCTATTTACAGCGGTGACACTATT	1026
341	QY	LeuHisIleGlnAspLysProcalaleuPheargSerGlyTyrIleTyrIleuLysProGly	360
1027	DB	CTGCACATCCACAGAACGACGCTCTATTCAGACATCTCTCAAGTGGCTTAAACGAGG	1086
361	QY	GlyLysValLeuIleSerAspTyrCysLysLysAlaglyProProSerProGluPheala	380
1087	DB	GGTAAGATTCTCATCTCACTGACTATTTCGAGAACTGCTGAACTCCGCTCTCTGAATTCGCA	1146
381	QY	AlaTyrIleLysGlnArgGlyTyrAspLeuHisAspValIleGlyIleTyrGlyGlnMetLeu	400
1147	DB	GAGTACATMAACAAGAGGATATGATCTACATGATGTTCAAGCTTACGACACAGATGCTG	1206
401	QY	LysAspalaGlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIleArgVal	420
1207	DB	MAAGCGCAGGCTTTGACGACGTTATTCGTCGAGGACCGTACTGATCAGTTTGTCAGATC	1266
421	QY	LeuArgLysGluLeuGluThrValGluLysGluLysAspValPheIleSerAspPheSer	440
1267	DB	CTCAGGCTGGAATTAGAAAAANGTGGAGAAGAAAAGGAAGAAATTCATCACGCACTCTCTCA	1326
441	QY	GluGluAspTyrAsnAspIleValGlyGlyTyrPasnAspLysLeuArgArgThrAlaLys	460
1327	DB	GAAGAGGATTACATGATGACATTGTCGAGATGTCGGCAAGCTTGAAGGCACTGCACTCT	1386
461	QY	GlyGluGlnArgTrpGlyLeuPheValAlaLysLysLys	473
1387	DB	GGTGACACAGAAATGGGATATTTCATATGACGCGACAGAAAG	1425
RESULT 7			
AF428454			
LOCUS		1656 bp mRNA linear PLN 18-OCT-2001	
DEFINITION		Arabisopsis thaliana At1948600/TLIN5_20 mRNA, complete cds.	
ACCESSION		AF428454	
VERSION		AF428454.1 GI:16226648	
KEYWORDS		PLI CDNA	
SOURCE		Arabisopsis thaliana (thale cress)	
ORGANISM		Arabisopsis thaliana	
REFERENCE		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
AUTHORS		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	
		rosids; eurosids II; Brassicales; Brassicaceae; Arabidoideae.	
		1 (bases 1 to 1656)	
		Chen, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Barb, J.,	
		Bowyer, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,	
		Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G.,	
		Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M.,	
		Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K.,	
		Quach, H.T., Sakurai, T., Samura, M., Seki, M., Southwick, A., Tang, C.C.,	
		Tohnumi, M., Takami, K., Yamamura, F., Yu, G., Yu, S., Shinohara, K.,	
		Davis, R.W., Theologis, A. and Eckert, D.B.	
		Arabisopsis cDNA clones	
JOURNAL		Unpublished	
REFERENCE		2 (bases 1 to 1656)	

85 AAAGAAGACGCTCTGAGTACTCTCTTTAATCCACCATATGTAAGGGAATCTGCTT 144  
 41 GtUleuGlyValAGlyIleGlyValGlyPheThrGlyGluLeuAlaGlyValGlyVal 60  
 145 GAACTTGGAGCTGGTATTGCTGCTTCACTGCTGATTCCTCAAAAGGCTGGTGAAT 204  
 61 IleAlaLeuAapPheIleGluSerAlaIleGlyValGluValIleAlaGlyHisTyr 80  
 205 ATCGCTCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 264  
 81 LysAenValLysPheMetCysAlaAapValThrSerProThrLysSerPheProHis 100  
 265 AAGAACATCAAGATTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 144  
 101 SerLeuAapValIlePheSerAsnTrpLeuMetTyrLeuSerAspGluValGlu 120  
 325 TCTATCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 384  
 121 AenLeuValGluArgMetLeuLysTrpLeuLysProGlyGlyTyrIlePhePheArgGlu 140  
 385 CTAAATGGCAGAGAGAAATGATGGATGGTCAAGCCAGGGGATACATTTCTTCAGAGA 444  
 141 SerCysPheHisGlnSerGlyAspHisLysLysSerAsnProThrHisTyrArgGlu 160  
 445 TCTTGTCTTCCATCAATCTGGGAGCAGCAGCAGGAAAGTCAAAACCCCACTCACTG 504  
 161 ProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAapGlySerGlyAsnSer 180  
 505 CCCAGTTCTACACAAAGTTTCCAGAAATGTCACACACGATGATGCTTCTGGCAATTC 564  
 181 TyrGluLeuSerLeuSerCysLysCysAlaGlyValLysValArgLysValAsn 200  
 565 TTTCAGCTCTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 624  
 201 GlnAsnGlnIleSerTrpLeuTrpGlnLysValAapSerLysAspAspLysGlyPheGln 220  
 625 CAGAATCAGATTCTCGATATGCAAAAGTCCAGCTGGAGATGACAGAGATTTCCAG 684  
 221 ArgPheLeuAapThrSerGlnTyrLysCysAenSerIleLeuArgTyrGluArgValPhe 240  
 685 GCTTCTTGGAATGTTCAATCAAGTCTAGTGGATCTGCTGCTGCTGCTGCTGCTGCT 744  
 241 GlyProGlyTyrValSerThrGlyTyrGluThrThrLysGluPheValSerMetLeu 260  
 745 GGGGAAGATATGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 804  
 261 AspLeuLysProGlyGlnLysValLeuAapValGlyCysGlyIleGlyGlyAspPhe 280  
 805 GACCTTAAACCGGACGAAAGTCTAGATGTTGGTGTGGTATCGGTGGAGGTGACTTC 864  
 281 TyrMetAlaGluThrPheAapValGluValGlyPheAapSerValAsnMetIle 300  
 865 TACATGCTGAGAAATTCGATGTTTCATGTTGGAATCGATCTGCTGCTGCTGCTGCTGCT 924  
 301 SerPheAlaLeuGluArgSerIleGlyLysLysCysAlaValGluPheValAlaAap 320  
 925 TCTTTCGCTGAGCGGGCCATGCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 984  
 321 CysThrLysIleLeuMetProAenSerPheAapValIleTyrSerArgAspThrIle 340  
 985 TGCACCAACCAACATATCCGATAATCTCTTGTATGCTCAATTCAGCGCTGACATATT 1044  
 341 LeuHisIleGlnAapLysProAlaLeuPheArgSerPheTyrLysTrpLeuLysProGly 360  
 1045 CTGCACATCCAGACAGCCAGCTCTATTTCAGACATCTTCATGCTGCTTAAACACAGG 1104  
 361 GlyLysValLeuLysSerAspTyrCysLysLysAlaGlyProProSerProGluPheAla 380  
 1105 GGTAAATCTCTCATCATGCTATTCAGAGCTGCTGAAATCCGCTCTCTGCTGCTGCT 1164  
 381 AlaTyrIleLysGlnArgGlyTyrAspLeuHisAspValLysGluTyrGlyGlnMetLeu 400

1165 GAGTACATATAAACAAGAGGATATGATCTACATGATGTTTCAAGCTTACGACAGATGCTG 1224  
 401 LysAapAlaGlyPheValAapValLeuAlaGluAapArgThrGluGlnPheIleArgVal 420  
 1225 AAAGACGAGCGCTTTCAGACGCTTATCGCTGAGGACGCTACTGATCAGTTGTTACAGTC 1284  
 421 LeuArgLysGluLeuGluThrValGluLysGluLysAapValPheIleSerAspPheSer 440  
 1285 CTCAGCGTGAATTAGAAAAAGTGGAGAAAGAAAGAAAGATTCATCAGCGACTTCTCA 1344  
 441 GtUleuAapTyrAenAapIleValGlyCysTrpAsnAapLysLeuArgArgThrAlaLys 460  
 1345 GAAGAGGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1404  
 461 GlyGluGlnArgTrpGlyLeuPheValAlaLysLysLys 473  
 1405 GGTGAACAGAAATGGGATTAATTCATAGCCGACAGAG 1443

# RESULT 8 AY063866

LOCUS 1660 bp mRNA linear PLN 18-SEP-2002  
 DEFINITION Arabidopsis thaliana putative phosphoethanolamine  
 N-methyltransferase (At1g98600) mRNA, complete cds.

ACCESSION AY063866  
 VERSION AY063866.1 GI:17390783

KEYWORDS FLI CDNA.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

# REFERENCE

## AUTHORS

Yanada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K.,  
 Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L.,  
 Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,  
 Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,  
 Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,  
 Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K.,  
 Davis, R.W., Ecker, J.R. and Theologis, A.  
 Arabidopsis Full Length cDNA Clones

Unpublished

# JOURNAL

## REFERENCE

## AUTHORS

Yanada, K., Banh, J., Banno, P., Chang, E., Dale, J.M., Goldsmith, A.D.,  
 Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C.,  
 Yamamura, Y., Yu, G., Yu, S., Bowser, L., Carninci, P., Chen, H.,  
 Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A.,  
 Karlin-Neumann, G., Kawai, J., Kim, C., Koesema, E., Lam, B., Lin, J.,  
 Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,  
 Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A.,  
 Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.  
 Direct Submission

# TITLE

## JOURNAL

## COMMENT

Submitted (19-NOV-2001) Plant Gene Expression Center, 800 Buchanan  
 Street, Albany, CA 94710, USA  
 RIKEN Genomic Sciences Center (GSC) members carried out the  
 collection and clustering of RAFL cDNAs (RAFL CDNA : 'RIKEN  
 Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J.,  
 Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,  
 Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGSC (SSP) Consortium members carried out the  
 sequencing and annotation of the RAFL cDNAs: Yamada, K., Banh, J.,  
 Banno, P., Chang, E., Dale, J.M., Goldsmith, A.D., Lee, J.M.,  
 Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C.,  
 Yamamura, Y., Yu, G., Yu, S., Bowser, L., Chen, H., Cheuk, R.,  
 Karlin-Neumann, G., Kawai, J., Kim, C., Koesema, E., Lam, B., Lin, J.,  
 Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,  
 Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A.,  
 Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamaoka, K. (SSP/PGSC) and Seki, M. (RIKEN GSC) contributed equally to  
 this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGSC)  
 contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis

genome submitted to GenBank.  
 FEATURES  
 Location/Qualifiers  
 1..1660  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /chromosome="1"  
 /clones="RAP106-86-D12 (R11999)"  
 /ecotype="Columbia"  
 /note="This clone is in a modified pBluescript vector (Lambda ZAP) as a XhoI/SstI insert."  
 1..1660  
 /gene="At1g48600"  
 1..34  
 /gene="At1g48600"  
 35..1462  
 /gene="At1g48600"  
 /codon\_start=1  
 /evidence=experimental  
 /product="putative phosphoethanolamine N-methyltransferase"  
 /protein\_id="AA136222.1"  
 /db\_xref="GI:17380784"  
 /translation="MEHSDLVTEAMMLDSKLDKBERPEVLSLIPPEKGVLEL  
 GAGRTPTGLAOKAGEVIALDFIESAIQKNSVNGHYKIKFMCADVTSPLDKIKD  
 SIDILFNNLLAYLSKDEKLEMAERIMGWKPGVIFPESCFHSGDSEKSNPHY  
 REPRFTYVQECOTRDLSCNSPELSMGCKICGAVNKNQNOICWIKQVND  
 KDFRFTLDVQKSSGILRYRVEGYSVTGSPRTPEFAKMDLKPQKVLVDGSG  
 IGGDFTMENFDVHVGGIDLSVNNISPALEAIGLKSVEFEVADCTTKYPRNSD  
 VIYSDTLIHQIDPFLPRTFFKWLAPGKGLVITDTCRAETPSPEFARYIKQGYD  
 HDVQYQGMKDGDFVIAEDRTDQVVLRRLEKVEKEEKFISFSEEDNDIV  
 GGWSAKLERTASGQKGLFIADKK"  
 misc\_difference 316  
 /gene="At1g48600"  
 /note="compared to genomic sequence"  
 /replace="t"  
 1463..1660  
 /gene="At1g48600"

Alignment Scores:  
 Pred. No.: 6,198-194 Length: 1660  
 Score: 2077.00 Matches: 379  
 Percent Similarity: 90.06% Conservative: 47  
 Best Local Similarity: 80.13% Mismatches: 47  
 Query Match: 82.58% Indels: 0  
 DB: Gaps: 0

US-10-031-331B-40 (1-473) x AY063866 (1-1660)

Qy 1 HisThrValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuApp 20  
 Db 41 CATCTAGTGAATTTGACTGTTGAAGCTATGATGCTTGACTCTAAAGCTTCTGATCTTGAC 100  
 Qy 21 LysGluGluArgProGluIleLeuSerMetLeuProLeuGluGlyLysCysLeuLeu 40  
 Db 101 AAGAGAGAGCTGCTGAGGTACTCTCTTAATCCACCATATGAAGGAAATCTGTGCTT 160  
 Qy 41 GluLeuGlyAlaGlyIleGlyArgPheThrGlyGluLeuAlaGlyLysAlaGlyGlnVal 60  
 Db 161 GAATCTGGAGCTGGTATGTTGCTTCACTGCTGATGCTTCAAAAGGCTGTGTGANGTT 220  
 Qy 61 IleAlaLeuAspPheIleGluSerAlaIleLysLysAsnGluValIleAsnGlyHisTyr 80  
 Db 221 ATCGCTCTTGACTTCATCCAAAGCGGATTCAGAAAGTGAAGTGTAAATGGGCATTAC 280  
 Qy 81 LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProHis 100  
 Db 281 AAGAACATCAAGTTATGTGTCTGATGTATACATCCACAGCTTGAAATCAAGATGGA 340  
 Qy 101 SerLeuAspValIlePheSerAsnTrpLeuLeuMetTyrLeuSerAspGluGluValGlu 120  
 Db 341 TCTATCGACTTGAATTTCTCAATCGTTGCTCATGTATCTCTGTATAAAGAGGTGGAA 400

Qy 121 AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyGlyTyrIlePhePheArgGlu 140  
 Db 401 CTAATGGCAGAGAAATGATGGATGGTCAAGCCAGGGGGATACATTTCTTTCAGAGAA 460  
 Qy 141 SerCysPheHisGlnSerGlyAspHisLysArgLysSerAsnProThrHisTyrArgGlu 160  
 Db 461 TCTTGCTTCCATCAATCTGGGACAGCAAGCCGAAAGTCAAAACCCCACTCACTACCGTGA 520  
 Qy 161 ProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSer 180  
 Db 521 CCCAGATTCTACAAAGGTTTTCAGGAATGTACAGACAGGTGCTTCTGGCAATTC 580  
 Qy 181 TyrGluLeuSerLeuLeuSerCysLysCysIleGlyAlaTyrValArgAsnLysLysAsn 200  
 Db 581 TTTGAGCTCTATGTTGGTTCGCAATTCATTTGGGGCTTATGTGAGAACACAGAGAAAT 640  
 Qy 201 GlnAsnGlnIleSerTrpLeuTrpGlnLysValAspSerLysAspLysGlyPheGln 220  
 Db 641 CAGATTCAGATTGCTGGATATGCAAAAGTCAAGCTGGAGATGACACAGGATTTCCAG 700  
 Qy 221 ArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArgValPhe 240  
 Db 701 CGTTTCTTGGCAATGTTCATTAAGTCTAGTGGGATCTTCGCTATGAGGCTGCTCTT 760  
 Qy 241 GlyProGlyTyrValSerThrGlyGlyTyrGluThrThrLysGluPheValSerMetLeu 260  
 Db 761 GGGGAAGGATATGTGACACTCGTGGATTTGACACCAACTAAAGAAATTTGTGGCAAGATG 820  
 Qy 261 AspLeuLysProGlyGlnLysValLeuAspValGlyCysGlyIleGlyGlyLysAspPhe 280  
 Db 821 GACCTTAACCGGACAGAAAGTCTAGATTTGGTGTGGTATCGTGGAGGTGCTTC 880  
 Qy 281 TyrMetAlaGluThrPheAspValGluValGlyPheAspLeuSerValAsnMetIle 300  
 Db 881 TACATGGCTGAGAAATTTGATGTTGATGTTGGAAATCGATCTGTGCGGTCAACATGATC 940  
 Qy 301 SerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluValAlaAsp 320  
 Db 941 TCTTTCGCACTGGAGCGGGCCATTTGACTCAATGCTCAGTTCAGTTCGAGTTCGAGTTC 1000  
 Qy 321 CysThrLysIleAsnTyrProAspAsnSerPheAspValIleTyrSerArgAspThrIle 340  
 Db 1001 TGCACCAACCAACATATCCGATATTCCTTTGATGTCATTTACAGCCGTGACACTATT 1060  
 Qy 341 LeuHisIleGlnAspLysProAlaLeuPheArgSerPheTyrLysTrpLeuLysProGly 360  
 Db 1061 CTGCACATCCAAAGCAAGCCAGCTCTATTCAAGGACATTTCTCAAGTGGGCTTAAACCCAGG 1120  
 Qy 361 GlyLysValLeuIleSerAspTyrCysLysLysAlaGlyProProSerProGluPheAla 380  
 Db 1121 GGTAAAGTTCTCATCTGACTATTTGCAAGAGTGTCAAACTCCGCTCTCTGAAATTCGCA 1180  
 Qy 381 AlaTyrIleLysGlnArgGlyTyrAspLeuHisAspValLysGluTyrGlyGlnMetLeu 400  
 Db 1181 GAGTACATAAAACAAAGAGGATATGATCTCATGATGTTTCAAGCTTACGACAGATGCTG 1240  
 Qy 401 LysAspAlaGlyPheValAspValIleAlaGluAspArgThrGluGlnPheIleArgVal 420  
 Db 1241 AAAGCGCAGGCTTTGACGAGCTTATCGTGGAGCCGCTACTGATCAGTTCGTTGTACAGTC 1300  
 Qy 421 LeuArgLysGluLeuGluThrValGluLysLeuLysAspValPheIleSerAspPheSer 440  
 Db 1301 CTGAGCGGTGAATTAAGAAAGTGGAGAAAGAAAGAGAGAAATTCATCAGGACTTCTCA 1360  
 Qy 441 GluGluAspTyrAsnAspIleValGlyTyrPheAsnAspLysLeuArgArgThrAlaLys 460  
 Db 1361 GAAGAGGATTAACATGACATTTGTTGGAGGATGCTCGCAAAAGCTTGAAGAGGTGATCT 1420  
 Qy 461 GlyGluGlnArgTrpGlyLeuPheValAlaLysLysLys 473  
 Db 1421 GGTGACAGAAATGGGATTTATCATGCCGACAGAG 1459



## RESULT 9

AF328858 1794 bp mRNA linear PLN 29-JAN-2001  
 Lycopersicon esculentum phosphoethanolamine N-methyltransferase  
 mRNA, complete cds.  
 AF328858 AF328858.1 GI:12584942  
 Lycopersicon esculentum (tomato)  
 Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
 1 (bases 1 to 1794)  
 Parani, M. and Parida, A.  
 Characterization of a cDNA for phosphoethanolamine  
 N-methyltransferase  
 Unpublished  
 2 (bases 1 to 1794)  
 Parani, M. and Parida, A.  
 Direct Submission  
 Submitted (13-DEC-2000) Plant Molecular Biology, M. S. Swaminathan  
 Research Foundation, III Cross Street, Chennai, Tamil Nadu 600 113,  
 India

## FEATURES

## source

1..1794  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4081"  
 106..11581  
 /codon\_start=1  
 /product="phosphoethanolamine N-methyltransferase"  
 /protein\_id="AA059894.1"  
 /db\_xref="GI:12584943"  
 /translations="MAAAGQREEVQKSYMIEHTAELTVEAMMLDSKAADLDKERRPE  
 VLSLPPEKTVLEAGAGIGFTGCELAGKAGQLALDPIEGAIKQENINIGHTKQVK  
 PMCADVTPNLIIPESVOLLIFSNLLATLSDEEVKALVERWVWLKVGHIIPRESC  
 PHOGDHKRNKTHTRDPSFTFKVFRCHYNAGDGKSFELSLAGCKICIGAYNKKQ  
 ONQICWTKQVTSKDMERFORLDTVOYKCSGLIERYVFOGQFVSGVLETKFEVA  
 MLDLRQKVLVDCGGGNGFNMAEKYDVHVVGIDLSINMISFALRAIGLCAVEF  
 EVADCTKTKTPDVTYSRDTLIHQDKPALFRPYKWLKRPQGVLLISDYCKRAGP  
 ASKEFEGYIKQRYDLDHVEAYQOMLDAGPHEVVAEDRTEQFIKVLQKSLDTVEKER  
 ESPTHEFSEQDYNEIVGGWKSLLIRSSGQERWGLFIKKK"

## ORIGIN

Alignment Scores:  
 Pred. No.: 1,38e-182 Length: 1794  
 Score: 2063.00 Matches: 379  
 Percent Similarity: 89.85% Conservative: 46  
 Best Local Similarity: 80.13% Mismatches: 48  
 Query Match: 82.03% Indels: 0  
 DB: 8 Gaps: 0

US-10-031-331B-40 (1-473) x AF328858 (1-1794)

QY 1 HisThrValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp 20  
 DB 160 CACACTGGCAACTCACTGTCGAGGCGATGCTCGACTTAAAGCAGCGGATCTCGAC 219  
 QY 21 LysGluGlnArgProGluIleLeuSerMetLeuProProLeuGluGlyIleCysLeuLeu 40  
 DB 220 AAAGAAGAGAGACCTGAGGTGCTCTCTGCTTCGCGCCGTATGAGGAAAAAATCTGTG 279  
 QY 41 GluLeuGlyAlaGlyIleGlyArgPheThrGlyGluLeuAlaGluLysAlaGlyGlnVal 60  
 DB 280 GAACCTGGCTGAGGATATGCGCCATTCATGCTGATGATAGCCAGGATGAGGACGCTT 339  
 QY 61 IleAlaLeuAspPheIleGluSerAlaIleLysLysAsnGluValIleAsnGlyIleTyr 80  
 DB 340 ATAGCTGGACTTATTTGAAGCGCAATTAAGCAAGATGAATTAATATGCGCACTAT 399  
 QY 81 LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerProProHis 100

DB 400 AGRATGTCAAGTTCATGTGCTGATGTGACTTCCCGAATTTGATTTTTTTCACCGAA 459  
 QY 101 SerLeuAspValIlePheSerAsnTrpLeuLeuMetTyrLeuSerAspGluValGlu 120  
 DB 460 TCGTGGACTTGAATATTTCTTAACGTGCTACTGATGTATCTTCTCATGAAGAGTTAAG 519  
 QY 121 AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyGlyIleTyrIlePhePheArgGlu 140  
 DB 520 GCTCTTGTTCAGACAAATGGTTATATATGTTGAAAGTCGTGGTGCACATATTTTCAGAG 579  
 QY 141 SerCysPheHisGlnSerGlyAspHisLysArgLysSerAsnProThrHisTyrArgGlu 160  
 DB 580 TCACTCTTCATCATCATCAGGACCAACAGCGAAGAACCAACCCACCATATTCAGAG 639  
 QY 161 ProArgPheTyrThrLysAlaPheLysGluCysHisLysLeuGlnAspGlySerGlyAsnSer 180  
 DB 640 CCTAGCTTTTATACAAAGGTTTTTTCAGAGAAATGTCAGTAAATGCTGGTGAATCA 699  
 QY 181 TyrGluLeuSerLeuSerCysLysCysValIleGlyAlaTyrValAsnLysValAsn 200  
 DB 700 TTTGAACTTTCTCTGCTGGTTCGAAGTGCATTTGAGCTTATGTGAGACAAAGAAAT 759  
 QY 201 GlnAsnGlnIleSerTrpLeuTrpGlnLysValAspSerLysAspAspLysGlyPheGln 220  
 DB 760 CAAATCAGATTTGCTGGACGTGGCAGAGGTTACTTCTTAAGGATGACATGGAATTCAG 819  
 QY 221 ArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuAlaTyrGluArgValPhe 240  
 DB 820 CGTTTCTTGGATCTGTTTCAGTACAAAGTGTAGTGGCATCTAGCTATGACAGGTCTTT 879  
 QY 241 GlyProGlyTyrValSerThrGlyGlyTyrGluThrThrLysGluPheValSerMetLeu 260  
 DB 880 GGACAGGTTTGTGTAGCAGGAGGACTTGAACATACAAAAGAAATTTGTAGCCATGTTG 939  
 QY 261 AspLeuLysProGlyGlnLysValLeuAspValGlyCysGlyIleGlyGlyAspPhe 280  
 DB 940 GATCTTCAGCGTGGCCAAAAGTCTTGTATGTTGCTGTGGCATTTGGAGGAGGTAACCTC 999  
 QY 281 TyrMetAlaGluThrPheAspValGluValValGlyPheAspLeuSerValAsnMetIle 300  
 DB 1000 TACATGCTGAAAGATGATGTTTCATGTTGTGGCATTTGAGTGTCTTATCAATGATA 1059  
 QY 301 SerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluValAlaAsp 320  
 DB 1060 TCCTTTGCTCTGAGCGTGTCTATGCTCTCAAGTGTGCTGTTGAATTTGAGGTTCGTAT 1119  
 QY 321 CysThrLysIleAsnTyrProAspAsnSerPheAspValIleTyrSerArgAspThrIle 340  
 DB 1120 TGCACCAAGAAAACATATCTGATGTGTACATTTGATGTATATACAGTTCGGGATCTATC 1179  
 QY 341 LeuHisIleGlnAspLysProAlaLeuPheAspSerPheTyrLysTrpLeuLysProGly 360  
 DB 1180 CTTCACTCAGGACCAACCTGCACCTTCATATCTTCTACAAAGTGGCTGAGACCAAGGA 1239  
 QY 361 GlyLysValLeuLysSerAspTyrCysLysLysAlaGlyProProSerProGluPheAla 380  
 DB 1240 GGCAAGTCTTCTTAAAGTGATTTACTGCAAAAGAGCTGCAACAGCATCAAAAGAAATTCGAA 1299  
 QY 381 AlaTyrIleLysGlnArgGlyTyrAspLeuHisAspValLysGluTyrGlyGlnMetLeu 400  
 DB 1300 GGTATATTAAAGCAAGGGGTATGATTTACATGATGTGAAAGTATGGGACATGCTC 1359  
 QY 401 LysAspNlaGlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIleArgVal 420  
 DB 1360 AGAGATCGGGTTTTTCATGAAGTTGTGGCCGAGGATGCTACTGAACAGTTTCATTAAGGTT 1419  
 QY 421 LeuArgLysGluLeuGluThrValGluLysGluLysAspValPheIleSerAspPheSer 440  
 DB 1420 CTCGAAAGGAGGTAGATATCTGTGGAGGAGGAAAGAGAAATCTTTATTCATGAGTCTCT 1479  
 QY 441 GluGluAspTyrAsnAspIleValGlyGlyTyrAsnAspLysLeuArgThrAlaLys 460  
 DB 1480 GAACAAGACTATTAATGAATAGTTGGAGGTGGAGGTGCAAGCTTAATCAGGAGTTCACTCT 1539





QY 420 ValLeuArgLysGluLeuThrValGluLysGluLysAspValPheIleSerAspPhe 439  
 Db 1508 GTCTTCAGAGGAACTGGAAGAGTTCGAGAGGAAAGGAGGAGTTCATCAAGACTTC 1567  
 QY 440 SerGluGluAspTyrAsnAspIleValGlyTyrPheAspLysLeuArgThrAla 459  
 Db 1568 ACTGAGGTGACTACAGCATATCTGTGGTGGGAGGCAAACTGGTCAGGACTGGT 1627  
 QY 460 LysGluGluGlnArgTyrGlyLeuPheValAlaLysLysLys 473  
 Db 1628 TCTGGGAGCAGAGATGGGCTGTGTTATTGCGAAGGAAG 1669

## RESULT 11

AY136372

LOCUS

DEFINITION

Arabidopsis thaliana phosphoethanolamine N-methyltransferase,

putative (Atig73600) mRNA, complete cds.

ACCESSION

AY136372

VERSION

AY136372.1

KEYWORDS

F11 CDNA.

SOURCE

Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE

1 (bases 1 to 1834)

Southwick, A., Nguyen, M., Tripp, M., Palm, C.J., Jones, T., Wu, T.,

Carninci, P., Chen, H., Cheuk, R., Chan, M.M., Chang, C.H., Dale, J.M.,

Deng, J.M., Hayashizaki, Y., Heuan, V.W., Lee, J.M., Ishida, J.,

Kamliya, A., Kawai, J., Kim, C.J., Narusaka, M., Quach, H.L., Sakurai, T.,

Satou, M., Seki, M., Shinn, P., Tang, C.C., Toroumi, M., Wallender, B.K.,

Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Shinozaki, K.,

Ecker, J., Theologis, A. and Davis, R.W.

Direct Submission

Submitted (25-JUL-2002)

DNA Sequencing and Technology Center,

Stanford University, 855 California Avenue, Palo Alto, CA 94304,

USA

e-mail for correspondence: arabseq@sequence.stanford.edu

COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the

collection and clustering of RAPL cDNAs (RAPL cDNA: "RIKEN

Arabidopsis Full-length cDNA"): Seki, M., Narusaka, M., Ishida, J.,

Satou, M., Kamliya, A., Sakurai, T., Carninci, P., Kawai, J.,

Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGSC (SSP) Consortium members carried out the

sequencing and annotation of the RAPL cDNAs: Southwick, A.,

Nguyen, M., Tripp, M., Palm, C.J., Jones, T., Wu, T., Chen, H., Cheuk, R.,

Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Heuan, V.W., Lee, J.M.,

Kim, C.J., Quach, H.L., Shinn, P., Tang, C.C., Toroumi, M.,

Wallender, B.K., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S.,

Ecker, J., Theologis, A. and Davis, R.W.

Southwick, A. (SSP/Stanford) and Seki, M. (RIKEN GSC) contributed

equally to this work. Shinozaki, K. (RIKEN GSC) and Davis, R.W.

(SSP/Stanford) contributed equally to this work as PIs.

FEATURES

Location/Qualifiers

1..1834

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/db\_xref="taxon:3702"

/chromosome="1"

/clone="RAPL09-26-M24 (R17822)"

/ecotype="Columbia"

/note="This clone is in pBluescript"

1..1834

/gene="Atig73600"

73..1203

/gene="Atig73600"

/codon\_start=1

/product="phosphoethanolamine N-methyltransferase,

putative"

/protein\_id="AA097038.1"

/db\_xref="GI:22531068"

/translation="MAHSHTNGAISPFSKDLCEEREIQNTYWKHSVGLSVEAMMLD

SKASLDKREPEITLAFPIETGTVLFEAGIGRFTTBLAKAGQVIADVFIISVVK

TKNENHGYKNVKFLCADVTSNNMFPNSMDLIISNLLMYLSQDEVEDLAKMLQN

TKVGYIFPFRSCPHQDNKRNYPNTHYREPKFTKLKPCHEMDNEDGNSVLSVS

CKTGAVVRNKNQNLCLWKQVSSNDNRGFORPLDNOYKSSGILRYRVPGRGFV

STGLETTEKPEVMDLKLKQKVLNKGCGIGGDEYKAEENFDVDVVGLDLSVNMISPA

LEHALGLKCSVBEFVADCTKKEYPDNTFDVYSRDLILHIQV"

## ORIGIN

Alignment Scores:

Pred. No.: 2,24e-178 Length: 1834

Score: 2018.00 Matches: 370

Percent Similarity: 84.94% Conservative: 53

Best Local Similarity: 74.30% Mismatches: 49

Query Match: 80.24% Indels: 26

DB: 8 Gaps: 1

US-10-031-331B-40 (1-473) x AY136372 (1-1834)

QY 1 HisThrValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp 20  
 Db 166 CATTGAGTGGGATTGAGTGTGAAGCTATGATCTTCAAAAGCTTCTGACCTCGAC 225  
 QY 21 LysGluGluArgProGluIleLeuSerMetLeuProLeuGluGluLysCysLeuLeu 40  
 Db 226 AAAGAAGAAGCTCTCGAGACTTGGTTCTTCCACCTATTGAGGACACACAGTCTA 285  
 QY 41 GluLeuGluAlaGlyIleGlyArgPheThrGlyGluLeuAlaGlyGlyGlnVal 60  
 Db 286 GAGTTTGGTGTGGAATTTGGTGTGTTTACTGATTAAGCTCAGAGCGCGCCAGTCC 345  
 QY 61 IleAlaLeuAspPheIleGluSerAlaIleLysLysAsnGluValIleAsnGlyHisTyr 80  
 Db 346 ATTGGGGTGAATTCATTGAAGTGTATCAAAAGATGAGAACATTAACGGTCACTAC 405  
 QY 81 LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProHis 100  
 Db 406 AAGAAGCTCAATTTCTGTGGCTGATGTCACATCAACCAATATGAATCTTCAATGAG 465  
 QY 101 SerLeuAspValIlePheSerAsnThrLeuLeuMetTyrLeuSerAspGluValGlu 120  
 Db 466 TCTATGGATCTGATATCTCAACTGGCTGAATGATATCTCTCTGATCAAGAGGTTGNA 525  
 QY 121 AsnLeuValGluArgMetLeuLysTyrLeuLysProGlyGlyTyrIlePhePheArgGlu 140  
 Db 526 GATTGGCGCAAAAGATGTTACAATGGACAAAGGTTGGCGGTATATTTCTTTCGGGAG 585  
 QY 141 SerCysPheHleGlnSerGlyAspHisLysArgLysSerAsnProThrHisTyrArgGlu 160  
 Db 586 TCTTGTTCATCTAGTCTGGTGTATACAGCGGAGTACAAACCCACACACTACCGTGA 645  
 QY 161 ProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSer 180  
 Db 646 CCTAAATTTTACCAAGAGCTTTTCAAGAATGCCATATGAATGACGAAGATGGCAATTCG 705  
 QY 181 TyrGluLeuSerLeuLeuSerCysLysCysIleGlyValaTyrValaTyrAsnLysAsn 200  
 Db 706 TATCAACTCTCTTGGTAGCTGTAAATGCATTCATTCAGGCTTATGATGAGAAACAAAGAAC 765  
 QY 201 GlnAsnGlnIleSerTrpLeuThrGlnLysValAspSerLysAspAspLysGlyPheGln 220  
 Db 766 CAGAACAGATATGCTGGCTTTGGCAGAAAGTCAGTCTGGATATGATAGAGGCTTCCAA 825  
 QY 221 ArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArgValPhe 240  
 Db 826 CGCTTCTTGACCAATGTCAGTATAGTCTAGTGTATCTTACCTATGAGCGTGTCTTT 885  
 QY 241 GlyProGlyTyrValSerThrGlyGlyTyrGluThrThrLysGluPheValSerMetLeu 260  
 Db 886 GGAGAGGGTTTGTGTAGCACAGGGGACCTCGAGACACAAAGAAATTCGTGGATATCTGT 945



QY 41 GluLeuGlyAlaGlyIleGlyAargpThrGlyGluLeuAlaGluLysAlaGlyGlnVal 60  
 Db 175 GAACCTGGAGCTGGTATTGGTCTGTTCTACCTGGTGAATAGCTCAAAAGGCTGGTGAATC 234  
 QY 61 IleAlaLeuAspPheIleGluSerAlaIleLysLysAsnGluValIleAsnGlyHisTyr 80  
 Db 235 ATTGCTCTGACTCTGATTTGATTAACCTGTTTCAAGAGAGTAAAGATATCAATGGGCAATAC 294  
 QY 81 LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProHis 100  
 Db 295 AAGAAATGTCAGGTTTATGTGCTGATGTTTACATCCCTGACCTCAAGATCACTGATGGA 354  
 QY 101 SerLeuAspValIlePheSerAsnTrpLeuMetTyrLeuSerAspGluGluValGlu 120  
 Db 355 TCTCTGACTGATTTCTCAACTGGCTGCTCATGTATCTTCTGCAAGAGGTGGAG 414  
 QY 121 AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyGlyTyrIlePhePheArgGlu 140  
 Db 415 CTTTGGCAGAAAGGATGGTGGTGGATCAAGGTTGGAGGATACATTTCTTCCTCGTGA 474  
 QY 141 SerCysPheHisGlnSerGlyAspHisLysAlaGlyLysSerAsnProThrHisTyrArgGlu 160  
 Db 475 TCTTCTTCCACCAATCAGGGACAGTAAAGCGGAATCCACCCCACTCACTACCGTGAA 534  
 QY 161 ProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSer 180  
 Db 535 CCCGCTTCTATCCCAAGGCTTTTCAAGAGTGTCAAGTCTGGGATGCTGCTGGGAATTC 594  
 QY 181 TyrGluLeuSerLeuSerCysLysCysLysGlyAlaTyrValArgAsnLysLysAsn 200  
 Db 595 TTTGAGCTCTATGATCGGATCGGATGCAAGTGCATTTGGAGCTTATGTCAAGAACCAAGAAAT 654  
 QY 201 GlnAsnGlnIleSerTrpLeuTrpGlnLysValAspSerLysAspAspLysGlyPheGln 220  
 Db 655 CAGAAATCAGATTTGTTGGATATGGCAGAGGTCAGCTCAGAAATGACAGAGGCTTCCAA 714  
 QY 221 ArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArgValPhe 240  
 Db 715 GCTTCTTGCAATGTCATACCAATCCAGTGGAAATCCTACGCTATGAGCGTGTCTTT 774  
 QY 241 GlyProGlyTyrValSerThrGlyTyrGlyTyrGluThrLysGluPheValSerMetLeu 260  
 Db 775 GGCCAGAGGTTTGTGAGCACTGTGCTGTGAGCAACCAAGAAATTTGTGGAGAAATG 834  
 QY 261 AspLeuLysProGlyGlnLysValLeuAspValGlyCysGlyIleGlyGlyAspPhe 280  
 Db 835 AATCTGAAACACGACAGCAAGATCTTAGATGTGGGTGGCATTTGGTGGAGGTGACTTC 894  
 QY 281 TyrMetAlaGluThrPheAspValGluValValGlyPheAspLeuSerValAsnMetIle 300  
 Db 895 TACATGGCTGAGAAATTTGATGTTTCACTGTTTGGTATCGATCTTTCTGTCAACATGATC 954  
 QY 301 SerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluValAlaAsp 320  
 Db 955 TCTTTCGATGGAAGCTGCTATTGGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1014  
 QY 321 CysThrLysIleAsnTyrProAsnSerPheAspValIleIleTyrSerArgAspThrIle 340  
 Db 1015 TGACACCAAAACACATACCAGATAATGTTTGTATGTCATTTACAGCCGCTGACATATT 1074  
 QY 341 LeuHisIleGlnAspLysProAlaLeuPheArgSerPheTyrLysTrpLeuLysProGly 360  
 Db 1075 CTGCACATCCACAGCAAAACAGCTGTTGTTAGGACTTCTTCAATGGCTTAACCCGGGA 1134  
 QY 361 GlyLysValLeuIleSerAspTyrCysLysLysAlaGlyProProSerProGluPheAla 380  
 Db 1135 GGTAAAGTCTCATCAGGACTACTGTAGAGCCCAAAATCCATCTCTGCTGAGTTTCA 1194  
 QY 381 AlaTyrIleLysGlnArgGlyTyrAspLeuHisAspValLysGluTyrGlyGlnMetLeu 400  
 Db 1195 GAGTACATCAACAGAGAGATATGATCTCCATGACGTTTCAAGCTTATGACAGATGCTA 1254

QY 401 LysAspAlaGlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIleArgVal 420  
 Db 1255 AAGACCTGCTCTCTACTGATGTGATCGGAGAGGACCGTACTGATCAGTTTATGCACTC 1314  
 QY 421 LeuArgLysGluLeuGluThrValGluLysGluLysAspValPheIleSerAspPheSer 440  
 Db 1315 CTGAACCTGAAATAGACAGGCTGGAGAAAGAAAGAAATTCATCTCCGACTCTCTCC 1374  
 QY 441 GluGluAspTyrAsnAspIleValGlyGlyTyrAsnAspLysLeuArgThrAlaLys 460  
 Db 1375 AAGAGGATACCATGATCATCTGTTGGAGGTGGAAGTCAAGCTGGAGAGGTGTGCTATC 1434  
 QY 461 GlyGluGlnArgTrpGlyLeuPheValAlaLysLys 472  
 Db 1435 GATGAGCAGAAATGGGAGCTTTTTCATCGCCAAACAG 1470

AV058175 1770 bp mRNA linear PLN 04-NOV-2001  
 Arabidopsis thaliana AT3g18000/MEB5\_22 mRNA, complete cds.

AV058175  
 DEFINITION Arabidopsis thaliana AT3g18000/MEB5\_22 mRNA, complete cds.  
 ACCESSION AV058175  
 VERSION AV058175.1  
 KEYWORDS GI:16648796  
 SOURCE FLI CDNA.  
 ORGANISM Arabidopsis thaliana (thale cress)

REFERENCE 1 (bases 1 to 1770)  
 AUTHORS Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J., Bower L., Carninci P., Date J.M., Goldsmith A.D., Hayaishizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A. and Becker J.R.  
 TITLE Arabidopsis cdna clones  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1770)  
 AUTHORS Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J., Bower L., Carninci P., Date J.M., Goldsmith A.D., Hayaishizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A. and Becker J.R.  
 TITLE Direct Submission  
 JOURNAL Submitted (28-AUG-2001) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL CDNA: 'RIKEN Arabidopsis Full-length CDNA'): Seki M., Narusaka M., Ishida J., Satou M., Kamiya A., Sakurai T., Carninci P., Kawai J., Hayaishizaki Y. and Shinozaki K.

The Salk, Stanford, RGC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Shinn P., Banh J., Bower L., Date J.M., Goldsmith A.D., Jiang P.X., Jones T., Karlin-Neumann G., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Southwick A., Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S., Davis R.W., Theologis A., and Becker J.R.

Cheuk R. (SSP/Salk) and Seki M. (RIKEN GSC) contributed equally to this work. Shinozaki K. (RIKEN GSC) and Becker J.R. (SSP/Salk) contributed equally to this work as PIs.

Location/Qualifiers  
 1. 1770

```

/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="3"
/cclone="RAFL09-10-N16 (R14028)"
/ecotype="Columbia"
1..215
216..1691
/notes="putative methyltransferase"
/codon_start=1
/product="AT3g18000/MEBS_22"
/protein_id="AAL25589.1"
/db_xref="GI:16648797"
/tranlations="MAAAYEERDIDKNYIWHSDLVTHAMMLDSRASLDKEERPE
VLCADTPYKSVLEKLCAGTGTGELAQKAGELALDPIIDNVIKNESINHYKNVK
PMADVSPDLKTDGSLDIFSNMLLVSLDKVELLABRMVGMVGVYIFRSC
PHOGDSKRKNSTHYRPFYKQFQSCOTDAGNSFELSMIGCKICIGAYVKNKN
ONQICHTWOKVSSNDRGFORFLDNVOKSSGILYRERFVGQFVSTGGLETKEFVE
KNMLKQKVLVGGCGIGGDFYMAEKEDVRVVGIDLSVNNISPALEAEGLSVPE
EVADCTTHKPDNSFDVYISROTILIHQIDKPAFRTFFKMLKPGKVLISDYCRSPK
SPAFSFIKQRYDLDHVOAYGQMLKAGFDVITARDTIDPQVNLKRELDNRVEKEK
EKFLSDPSKEDYDDIDVGGWKSLERCAQSDQKWLFIANKN"
1692..1770

```

## 3' UTR

ORIGIN

## Alignment Scores:

```

Pred. No.: 9,61e-178 Length: 1770
Score: 2011.00 Matches: 367
Percent Similarity: 98.56% Conservative: 51
Best Local Similarity: 77.75% Mismatches: 54
Query Match: 79.96% Indels: 0
DB: 8 Gaps: 0

```

US-10-031-331B-40 (1-473) x AY058175 (1-1770)

```

QY 1 HisThrValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp 20
DB 270 CATTCCGCTGATCTGACTGTGTAAGCTATGATGCTTGTGACTCGAGAGCTTCTGATCTCGAC 329
QY 21 LysGluGluArgProGluIleLeuSerMetLeuProProLeuGluGluLysCysLeuLeu 40
DB 330 AAGGAAGAACCTCTGAGGTACTCTCTTGTCTCCCTCCATATGAAGGCAATTCAGGTGTG 389
QY 41 GluLeuGlyAlaGlyIleGlyArgPheThrGlyGluLeuAlaGluAlaGlyGlnVal 60
DB 390 GAACCTGAGCTGGTATTGCTGCTTCTGCTGTAATAGCTCAAAAGGCTGCTGAACCTC 449
QY 61 IleAlaLeuAspPheIleGluSerAlaIleLysLysGluValIleAsnGlyHisThr 80
DB 450 ATTCCTCTGACTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 509
QY 81 LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProPheHis 100
DB 510 AAGAATGTCAAGATTTATGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 569
QY 101 SerLeuAspValIlePheSerAsnTrpLeuLeuMetTyrlleuSerAspGluValGlu 120
DB 570 TCTCTGACTGATTTCTCCAACTGCTGCTCATGATCATGATCATGATCATGATCATGATCATG 629
QY 121 AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyTyrllePhePheArgGlu 140
DB 630 CTTTGTGGCAGAAAGATGCTCGGTGATCAAGGTGGAGGATACATTTCTTCTCGGTGAA 689
QY 141 SerCysPheHisGlnSerGlyAspHisIlyeArgLysSerAsnProThrHisTyrlleGlu 160
DB 690 TCTTGTCTCCACCAATCAGGCGACAGTAAAGCGGAATCCAAACCCCACTCACTACCGTGA 749
QY 161 ProArgPheTyrllePheAlaPheLysGluCysHisIleuLeuAspGlySerGlyAsnSer 180
DB 750 CCGCGTTCATCTCAGGTCTTCTCAGGTGTCAGACTCGGATGCTGCTGGAATATCA 809
QY 181 TyrGluLeuSerLeuLeuSerCysLysCysIleGlyAlaTyrlleValArgAsnLysLysAsn 200
DB 181 TyrGluLeuSerLeuLeuSerCysLysCysIleGlyAlaTyrlleValArgAsnLysLysAsn 200

```

```

DB 810 TTTGAGCTCTCTATGATCGGATCGAAGTCGATTGGAGCTTATGTCACAGAACAGAGAAAT 869
QY 201 GlnAsnGlnIleSerTrpLeuTrpGlnLysValAspSerLysAspAspLysGlyPheGln 220
DB 870 CAGAAATCAGATTTGTTGGATATGCGCAGAGAGTTCAGCTCAGAAAATGACAGAGGCTTCCAA 929
QY 221 ArgPheLeuAspThrSerGlnTyrlleCysCysAsnSerIleLeuArgTyrlleGluArgValPhe 240
DB 930 CGTTTCTTGGCAATGTCCAAATACAAATCCATCGAATCTTACGCTATGAGGGTGTCTTT 989
QY 241 GlyProGlyTyrlleValSerThrGlyTyrlleGluThrThrLysGluPheValSerMetLeu 260
DB 990 GGCCAAAGGGTTTGTGAGCACTGTGTGAGTTCAGACCAACCAAGAAATTTGTGGAGAAATG 1049
QY 261 AspLeuLysProGlyGlnLysValLeuAspValGlyCysGlyIleGlyGlyAspPhe 280
DB 1050 AATCTGAAACAGCAGACAGAAAGTCTTGTAGATGTTGGGTGTGGCATTTGGTGGAGGTGACTTC 1109
QY 281 TyrlleAlaGluThrPheAspValGluValGlyPheAspLeuSerValAsnMetIle 300
DB 1110 TACATGCTGAGAGTTGATGTTACGTTGTGTGATGATGATGATGATGATGATGATGATGATG 1169
QY 301 SerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluValAlaAsp 320
DB 1170 TCTTTCGCAATGGAACTGCTATTTGGACTCAGCTGCTCGTTGAGTTTGGAGTTTGGCTGAT 1229
QY 321 CysThrLysIleAsnTyrlleProAspAsnSerPheAspValIleTyrlleSerArgAspThrIle 340
DB 1230 TGCACCAACAAACACTACCCAGATATTCGTTGTGATGATGATGATGATGATGATGATGATG 1289
QY 341 LeuHisIleGlnAspLysProAlaLeuPheArgSerPheTyrlleTyrlleLysProGly 360
DB 1290 CTGCACATCTCCAAAGCAACAGCCTTGTGTTAGGACTTCTTCAATGAGCTTAAACCGGGA 1349
QY 361 GlyLysValLeuIleSerAspTyrlleCysLysAlaGlyProProSerProGluPheAla 380
DB 1350 GGTAAAGTCTTCATCAGCAGACTACTGTAGAAGCCCCAAAACCTCCTGCTGAGTTTCA 1409
QY 381 AlaTyrlleLysGlnArgGlyTyrlleAspLeuHleAspValLysGluTyrlleGlyGlnMetLeu 400
DB 1410 GAGTACATCAACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1469
QY 401 LysAspAlaGlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIleArgVal 420
DB 1470 AAGAAGCTGCTCTTCACTGATGTCGAGAGGACCGTACTGATGATGATGATGATGATGATG 1529
QY 421 LeuArgLysGluLeuGluThrValGluLysGluLysAspValPheIleSerAspPheSer 440
DB 1530 CTGAAACCTGAAATAGACAGGTGGAGAAAGAAAGAAAGAAATTCATCTCCGACTTCTCC 1589
QY 441 GluGluAspTyrlleAsnAspIleValGlyTyrlleAsnAspIleValArgThrAlaLys 460
DB 1590 AAGAAGGATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1649
QY 461 GlyGluGlnArgTrpGlyLeuPheValAlaLysLys 472
DB 1650 GATGAGCAGAAATGGGACTTTTTCATCGCCCAACAG 1685

```

## RESULT 14

```

AF367299 1809 bp mRNA linear PLN 07-FEB-2002
LOCUS Arabidopsis thaliana AT3g18000/MEBS_22 mRNA, complete cds.
DEFINITION AF367299.1 GI:13605800
ACCESSION AF367299.1
VERSION FLI CDNA.
KEYWORDS Arabidopsis thaliana (chale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1809)
REFERENCE
AUTHORS Cheuk,R., Chen,H., Kim,C.J., Meyers,M.C., Shinn,P., Banh,J.,
Bower,L., Carninci,P., Chung,M.K., Goldsmith,A.D., Hayashizaki,Y.,

```











GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2004, 04:53:44 ; Search time 128 Seconds  
(without alignments)  
6945.557 Million cell updates/sec

Title: US-10-031-331b-39

Perfect score: 1602  
Sequence: 1 cacaccgtgtattaccat.....aaaaaaaaaaaaaaaaaaaa 1602

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 segs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA.\*  
1: /cgm2\_6/ptodata/2/ina/5A COMB.seq.\*  
2: /cgm2\_6/ptodata/2/ina/5B COMB.seq.\*  
3: /cgm2\_6/ptodata/2/ina/6A COMB.seq.\*  
4: /cgm2\_6/ptodata/2/ina/6B COMB.seq.\*  
5: /cgm2\_6/ptodata/2/ina/PCTUS COMB.seq.\*  
6: /cgm2\_6/ptodata/2/ina/backfileseq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	149.4	9.3	280	4	US-09-313-294A-6753
2	149	9.3	286	4	US-09-313-294A-7492
3	125.8	7.9	274	4	US-09-313-294A-3443
4	109.2	6.8	260	4	US-09-313-294A-876
5	101	6.3	174	4	US-09-313-294A-5689
6	100.4	6.3	274	4	US-09-313-294A-1188
7	57	3.6	171	4	US-09-313-294A-3925
8	46.6	2.9	7218	1	US-08-232-463-14
9	42.8	2.7	505	4	US-09-621-976-15639
10	42.6	2.7	1106	4	US-09-257-179-19
11	42.4	2.6	5789	3	US-09-242-548-3
12	41.6	2.6	1173	4	US-09-489-039A-2534
13	41.2	2.6	3927	4	US-09-293-238B-1
14	41.2	2.6	3933	3	US-08-199-776-1
15	41.2	2.6	3933	3	US-08-663-731-1
16	41.2	2.6	3933	3	US-08-879-338-1
17	41.2	2.6	3933	5	PCT-US95-02044-1
18	41.2	2.6	1664976	4	US-08-916-421B-1
19	40.6	2.5	1101	4	US-09-252-991A-14068
20	40.6	2.5	1542	4	US-09-252-991A-14202
21	40.4	2.5	1732	1	US-08-480-346-2
22	40.4	2.5	1732	2	US-08-243-541-2
23	40.4	2.5	1732	2	US-08-480-344-2
24	40.2	2.5	2058	2	US-08-749-391-1
25	40.2	2.5	2058	3	US-09-390-200-1
26	39.8	2.5	586	3	US-09-385-982-139
27	39.4	2.5	5852	1	US-07-867-106-2

28	39.2	2.4	946	3	US-08-916-443A-15	Sequence 15, Appl
29	39	2.4	891	4	US-09-134-001C-483	Sequence 483, App
30	39	2.4	1586	4	US-09-178-093B-3	Sequence 3, Appl
31	38.8	2.4	1342	4	US-09-489-847-89	Sequence 89, Appl
32	38.8	2.4	2170	4	US-09-807-258-9	Sequence 9, Appl
33	38.8	2.4	1230025	4	US-09-198-452A-1	Sequence 1, Appl
C 34	38.6	2.4	10968	2	US-08-680-327-2	Sequence 2, Appl
C 35	38.6	2.4	10968	3	US-09-228-246-1	Sequence 1, Appl
36	38.4	2.4	2128	4	US-09-280-116-11	Sequence 11, Appl
37	38.4	2.4	2459	4	US-09-443-795-2	Sequence 2, Appl
38	38.4	2.4	2745	4	US-09-833-381-1410	Sequence 1410, Ap
39	38	2.4	1224	4	US-09-543-681A-262	Sequence 262, App
C 40	37.8	2.4	1164	4	US-09-601-198-65	Sequence 65, Appl
41	37.2	2.3	195	4	US-09-621-976-15314	Sequence 15314, A
42	37.2	2.3	1064	1	US-08-378-588-15	Sequence 15, Appl
43	37.2	2.3	1064	2	US-08-811-094-15	Sequence 15, Appl
44	37.2	2.3	1064	5	PCT-US94-11121-15	Sequence 15, Appl
45	37.2	2.3	1270	1	US-08-378-588-23	Sequence 23, Appl

## ALIGNMENTS

## RESULT 1

US-09-313-294A-6753  
; Sequence 6753, Application US/09313294A  
; Patent No. 6476212

GENERAL INFORMATION:

APPLICANT: Lalgudi, Raghunath V.

APPLICANT: Ito, Laura Y.

APPLICANT: Sherman, Bradley K.

TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR

FILE REFERENCE: PL-0017 US

CURRENT APPLICATION NUMBER: US/09/313,294A

CURRENT FILING DATE: 1999-05-14

NUMBER OF SEQ ID NOS: 7600

SOFTWARE: PERL Program

SEQ ID NO 6753

LENGTH: 280

TYPE: DNA

ORGANISM: Zea mays

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: Incyte ID No. 6476212 700352341HI

US-09-313-294A-6753

Query Match 9.3%; Score 149.4; DB 4; Length 280;  
Best Local Similarity 71.0%; Pred. No. 5.4e-32;  
Matches 198; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY	1075	CCAGGAGCTAAAGTTCTAATCAGTATTAAGTCTCATGATGTAAAGGAATATGGCAG	1134
DB	2	CTGGGGCAGGTCCTTATCAGTACTCTGCGAGGTCCTGGGAAACCATCAGAGGA	61
QY	1135	TTCCGCGCTTACATTAAGCAGAGGGATATGATCTCATGATGTAAAGGAATATGGCAG	1194
DB	62	TTTCAGCGGTACATTAAGCAGAGGTTATGACATCATCTCTGTGGAGGCTTATGGACAG	121
QY	1195	ATGCTTAAAGATGCTGGATTTGTGATGTTCTTCCGAGATAGAACTGAGCAGTTCAAT	1254
DB	122	ATGTTGAAGAGTCTGGTTTCGTGATGTCATTCGTGAGATCGAATCATGTTCTT	181
QY	1255	CGAGTTCTACGAGGAGTACTAGAGCTGTGAGAGGAAAGATGTGTTCAATGAT	1314
DB	182	GTGTTTGTAGTAAGGAGTGTAGCTGAATTTGAAAGAACAGAGGAGATTTCTGTCTGAC	241
QY	1315	TTCTCTGAGGAGTATCAATGACATTTGTGGAGGTGG	1353
DB	242	TTCCAGGAGGAGTACTAGATGATATCTGTTGACGTTGG	280

## RESULT 2

US-09-313-294A-7492

Sequence 7492, Application US/09313294A  
Patent No. 6476212  
GENERAL INFORMATION:  
APPLICANT: Laligudi, Raghunath V.  
APPLICANT: Ito, Laura Y.  
APPLICANT: Sherman, Bradley K.  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR  
FILE REFERENCE: PL-0017 US  
CURRENT APPLICATION NUMBER: US/09/313,294A  
CURRENT FILING DATE: 1999-05-14  
NUMBER OF SEQ ID NOS: 7600  
SOFTWARE: PERL Program  
SEQ ID NO 7492  
LENGTH: 286  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: misc.feature  
OTHER INFORMATION: Incyte ID No. 6476212 700381824H1  
US-09-313-294A-7492

Query Match 9.3%; Score 149; DB 4; Length 286;  
Best Local Similarity 70.2%; Pred. No. 7.1e-32;  
Matches 200; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 400 GGTACATTTCTTCAGAGATCTGTTTCATCAATCTGGGATCACAACGCAAAAGC 459  
DB 2 GGCATATTTCTTTAGAGATCATGTTTTCCCAATCTGGAGTTCCAAAGGAGGTG 61  
QY 460 AATCCACCCATCTACCGTAACCTAGGTTCTACACTAAGCCCTTCAAGAGTGTCTATTG 519  
DB 62 AACCCACACATCTATCGAAGCAAGGTTTATACCAAGGTATTTAAAGGGCCATTCA 121  
QY 520 CAAGATGATCTGGAACCTCTTATGAGCTCTCCCTACTAGCTGCAAAATGATTGAGCT 579  
DB 122 TTGTATCAAGTGGAGGTGCTTGTGAACTTTCTCTAGTACCTGTAAATGATTGGGCT 181  
QY 580 TATGTCAGAACAGAAAACCAAGACAGATTAGTTGTTGGCAAAAGTTGATTTCT 639  
DB 182 TATGTCAAAACAGAGAGATCAAAACAGATATGCTGTTATGGAAGGTAATCA 241  
QY 640 AAGGATGATAAGGGTTCCAGCGATTTCTGGATCTAGCCAGTAC 684  
DB 242 ACAGAGACAGAGATTTCAAAGATTTCTGGACACAGCGCAATAC 286

RESULT 3  
US-09-313-294A-3443  
Sequence 3443, Application US/09313294A  
Patent No. 6476212  
GENERAL INFORMATION:  
APPLICANT: Laligudi, Raghunath V.  
APPLICANT: Ito, Laura Y.  
APPLICANT: Sherman, Bradley K.  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR  
FILE REFERENCE: PL-0017 US  
CURRENT APPLICATION NUMBER: US/09/313,294A  
CURRENT FILING DATE: 1999-05-14  
NUMBER OF SEQ ID NOS: 7600  
SOFTWARE: PERL Program  
SEQ ID NO 3443  
LENGTH: 274  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: misc.feature  
OTHER INFORMATION: Incyte ID No. 6476212 700611594H1  
NAME/KEY: unsure  
LOCATION: 52, 202  
OTHER INFORMATION: A, J, C, G, or other  
US-09-313-294A-3443

Query Match 7.9%; Score 125.8; DB 4; Length 274;

Best Local Similarity 70.6%; Pred. No. 2e-25;  
Matches 180; Conservative 0; Mismatches 74; Indels 1; Gaps 1;  
QY 655 TTCCAGCGATTTCTGGATATCTAGCCAGTACAGTGTAATAGCAATCTCGGATATGAGCGT 714  
DB 5 TTCAAAGATTTCTTGACACAGCTGCAATACAAAACAGTGGATATTTACGTTACGAGCGT 64  
QY 715 GTATTTGCCCTGTTATGTTAGCACTGGAGGATATCAACCAACCAAGAGTTCTGTCA 774  
DB 65 GTCTTTGGTGAAGGTTTGTGAGCCTCTGGAATCGACACCAAGAGATTTGTGGC 124  
QY 775 ATGCTGGACTTGAAGCCCTGCGCAAGCTCTCGATGTTGGTTGTGGAATTTGGTGAAGT 834  
DB 125 ATGCTCGATCTTAAACCGGCGCAGAAAGTACTTTGATGTCGATGTGGAATTCGAGCGCGC 184  
QY 835 GACTTTTACATGGCGGAGACCTTTGATGT-TGAGGTTGTTGGATTGATCTCTCCGTTAA 893  
DB 185 GACTTTTACATGGCTGCNACTATGATGTCTCCATGTTCTTGTATTGATCTATCGTGAA 244  
QY 894 TATGATTTCTTTTC 908  
DB 245 CATGGTTTCATTTGC 259

RESULT 4  
US-09-313-294A-876  
Sequence 876, Application US/09313294A  
Patent No. 6476212  
GENERAL INFORMATION:  
APPLICANT: Laligudi, Raghunath V.  
APPLICANT: Ito, Laura Y.  
APPLICANT: Sherman, Bradley K.  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR  
FILE REFERENCE: PL-0017 US  
CURRENT APPLICATION NUMBER: US/09/313,294A  
CURRENT FILING DATE: 1999-05-14  
NUMBER OF SEQ ID NOS: 7600  
SOFTWARE: PERL Program  
SEQ ID NO 876  
LENGTH: 260  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: misc.feature  
OTHER INFORMATION: Incyte ID No. 6476212 700549985H1  
NAME/KEY: unsure  
LOCATION: 24, 86  
OTHER INFORMATION: a, t, c, g, or other  
US-09-313-294A-876

Query Match 6.8%; Score 109.2; DB 4; Length 260;  
Best Local Similarity 68.5%; Pred. No. 7.8e-21;  
Matches 178; Conservative 0; Mismatches 80; Indels 2; Gaps 2;  
QY 599 ACCAGAACAGATTTAGTTGGTT-GTGGCAAAAGTTGATTTAAGGATGATAAGGGGTTTC 657  
DB 1 ATCAAAACAGATATGCTGTTANTGGGAAAGGTAATAATCAACAGAGACAGAGATTT 60  
QY 658 CAGGATTTCTGGA-TACTAGCCAGTACAAGTGAATAGCAATCTCGATATGAGCGCT 716  
DB 61 CAAGATCTCTGACACAGCTGCAATNCAAAACAGTGGGATATTTGTTACGAGCGTGT 120  
QY 717 ATTTGCCCTGTTATGTTAGCACTGGAGGATATGAACCAACCAAGAGTTTGTGCAAT 776  
DB 121 CTTTGTGGAAGGTTTGTGAGCCTGTTGGAATCGACACCAACAAAGGAATTTGTGGCGAT 180  
QY 777 GCTGGACTTGAAGCCCTGGCGCAGAGCTCTCGATGTTGGTTGGTAATTTGGTGAAGTGA 836  
DB 181 GCTGATCTTAAACCGGCGCAGAAAGTACTTTGATGTCGATGTGGAATTCGAGCGCGA 240  
QY 837 CTTTATCATGTCGAGACT 856  
DB 241 CTTTATCATGTCGCAACT 260

```
RESULT 5
US-09-313-294A-5689
; Sequence 5689, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 5689
; LENGTH: 174
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700350535H1
US-09-313-294A-5689

Query Match
Best Local Similarity 6.3%; Score 101; DB 4; Length 174;
Matches 128; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 701 TCGGATATGAGCGGTGATTTGGCCCTGGTATGTTAGCACTGGAGGATATGAACACCA 760
DB 2 TACGTTAGAGCGGTGCTTTTGGTGAAGGTTTGTGAGCACTGGTGAATCGAGACACNA 61
QY 761 AAGAGTTTGTGTCATGCTGACCTTGAAGCTGCGACAGGTCCTGGAATGTTGGTGTG 820
DB 62 AGGAATTTGCGGATGCTGATCTTAAACCGGCGAGAGTACTTGATGTCGATGTG 121
QY 821 GAATGTTGGAGGTGACCTTTACATGCGGAGACCTTTGATGTTGAGGTGT 873
DB 122 GAATTTGGAGGCGGACCTTTTACATGCTGCAAAATGATGATGTCATGTTCTT 174

RESULT 6
US-09-313-294A-1188
; Sequence 1188, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 1188
; LENGTH: 274
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700550521H1
; LOCATION: 256
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-1188

Query Match
Best Local Similarity...67.1%; Pred.No. 2.3e-18;
Matches 186; Conservative 0; Mismatches 87; Indels 4; Gaps 3;

QY 142 CGTTTACTGATGATGGCTGAGAAAGCTGGCCAGGTTATGCTCTCGATTTCAATTGAG 201
```

```
DB 2 CCCTTTACTGGAGATCTGGCAAAAGAGCTGGGCACGTTCTGGCTCTAGACTTCAATTGAA 61
QY 202 AGTGCTATCAAGAAGATGAATGAATCAATGGGCACCTACAAAATGTCTAGTTTATGTGT 261
DB 62 AGTGTGATTAAAGAAGAACCAAGCATTAATGGGCATCAACAAGAACATAACCTTCAGGTGT 121
QY 262 GCTGATGTGACTTCTCCCACTCTCCAGTTTCCCAACCATTCATTTGGATGTGATTTCTCC 321
DB 122 GCTGATGTGACATCTAAGCATTTGAAGATTAATA-ACTCTGTGATGATTTCA 180
QY 322 AATGTTTACTCATGATCTCTTCATGAAGAGGTGGAAAATTTGGTTCAAGAAGATTTG 381
DB 181 AACTGGCTATTAAATGTA-CTTTCAGATGAGAGGTCCAAAAGCTTGTGGGGAAATGGT-- 237
QY 382 AATGTTTGAAGCCAGCGGGGTACATTTTCTTCAGAG 418
DB 238 AATGCTTAAAGTCGGANGCCATATTTCTTTAGAG 274

RESULT 7
US-09-313-294A-3925
; Sequence 3925, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 3925
; LENGTH: 171
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700282372H2
; LOCATION: 5, 18, 60, 68, 104, 145, 163, 170
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-3925

Query Match
Best Local Similarity 63.2%; Pred.No. 2.1e-06;
Matches 98; Conservative 0; Mismatches 56; Indels 1; Gaps 1;

QY 1 CACACCGTTGATTTAAACCATTTGAAGCTGATGATGCTGATTCATCAAGCTTCTGATCTTGAC 60
DB 15 CACNCAAGTGCCTCACTGTCGAGTCATGATGCTGATTCATGATCCGTCGCGCA-NTGAC 73
QY 61 AAGAGAGAACTCTGAGATCTTTTCATGCTTCCGCTCTTTGAAGGAAATGCTCTTG 120
DB 74 AAGGAAGAGCGACCGAGATCCTGTCTTTGTTCCCTCTTACAAAGGGAATCAGTTCTA 133
QY 121 GAATCTGGGCTGTGTTTGTGTTTACTGTTGTA 155
DB 134 GAATCTGGTGGAGGATAGGACGCTTTACNGGAGA 168

RESULT 8
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, P. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
```

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Foley & Lardner  
 STREET: 1800 Diagonal Road, Suite 500  
 CITY: Alexandria  
 STATE: VA  
 COUNTRY: USA  
 ZIP: 22313-0299  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/232,463  
 FILING DATE:  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/07/935,313  
 FILING DATE:  
 APPLICATION NUMBER: EP 91 114 300.6  
 FILING DATE: 26-AUG-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: BENT, Stephen A.  
 REGISTRATION NUMBER: 29,768  
 REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703)836-9300  
 TELEFAX: (703)683-4109  
 TELEX: 899149  
 INFORMATION FOR SEQ ID NO: 14:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 7218 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 CLONE: pTZ9pt-Fls  
 US-08-232-463-14

Query Match 2.9%; Score 46.6; DB 1; Length 7218;  
 Best Local Similarity 3.3%; Pred. No. 0.011;  
 Matches 13; Conservative 219; Mismatches 163; Indels 0; Gaps 0;  
 QY 1051 AGATCTCTTACAAATGTTGAAGCAGGAGTAAAGTCTTATCAGTGAATCTGCAAG 1110  
 DB 1446 AGAATTGTCACRR 1387  
 QY 1111 AAAGCTGGTCCACCTCAGCTGATCCGCTTACATTAAGCAGGAGGATATGATCTC 1170  
 DB 1386 RRR 1327  
 QY 1171 CATGATGAAGGAATATGCGCAGATCTTAAAGATGCTGATTTGTTGATGTTCTGCC 1230  
 DB 1326 RRR 1267  
 QY 1231 GAGGATAGACTGAGCAGTTTCTGAGTCTTACGAGGAGTCTGAGAGCTGTTGAGAAG 1290  
 DB 1266 RRR 1207  
 QY 1291 GAAAGAGATGTTCTATTAGTGAATTTCTGAGGAGGATTAACAATGATGTTGAGGT 1350  
 DB 1206 RRR 1147  
 QY 1351 TCGAATGATGATGTTGCGGAGGACTGCGAAGGCTGAGCAACGATGCGGCTGTTGTTGCC 1410  
 DB 1146 RRR 1087  
 QY 1411 AAGAGAGATGAGATCAGTTCCGCACTGGCAC 1445  
 DB 1086 RRR 1052

RESULT 9

US-09-621-976-15639  
 : Sequence 15639, Application US/09621976  
 : Patent No. 6639063  
 : GENERAL INFORMATION:  
 : APPLICANT: Dumas Milne Edwards, J.B.  
 : APPLICANT: Giordano, J.Y.  
 : TITLE OF INVENTION: EST8 and Encoded Human Proteins.  
 : FILE REFERENCE: GENSET.054PR2  
 : CURRENT APPLICATION NUMBER: US/09/621,976  
 : CURRENT FILING DATE: 2000-07-21  
 : NUMBER OF SEQ ID NOS: 19335  
 : SOFTWARE: Patent.pm  
 : SEQ ID NO 15639  
 : LENGTH: 505  
 : TYPE: DNA  
 : ORGANISM: Homo sapiens  
 : US-09-621-976-15639  
 Query Match 2.7%; Score 42.8; DB 4; Length 505;  
 Best Local Similarity 17.8%; Pred. No. 0.032;  
 Matches 67; Conservative 144; Mismatches 161; Indels 4; Gaps 1;  
 QY 741 TCGAGGATATGAAACACCAAGAGATTGTGTCAATGCTGGACTTGAAGCCTGGCGAGAA 800  
 DB 107 YSGMMSKRWKRWKRWKRWKRWKRWKRWKRWKRWKRWKRWKRWKRWKRWKRWKRWK 166  
 QY 801 GGTCTCTGATCTGTGTGTGATTTGATGAGGTGAGTCTTTACATGCGCGAGACCTTTGA 860  
 DB 167 YMTYRTYMTYRTYMTYRTYMTYRTYMTYRTYMTYRTYMTYRTYMTYRTYMTY 226  
 QY 861 TGTGAGGTGTTGAGATTGATCTCTCGCTTAATATGATTTCTTGGCCCTTGAGGCTTC 920  
 DB 227 CYTMMCTYTTMCTCTYKKSYYTCTKSYTCTKSYTCTKSYTCTKSYTCTKSYTCT 286  
 QY 921 TATTGGCTTAAATGCTGTTGATTTGAGGTGAGGATTTGACCAAGATAAATACCTACCC 980  
 DB 287 CSKWKRWKRWKRWKRWKRWKRWKRWKRWKRWKRWKRWKRWKRWKRWKRWKRWK 346  
 QY 981 TGAATAA---CTCTTTGATGATCTATAGCCGTGACACCTTCTGATATTCAGGACA 1036  
 DB 347 CSARSAKGRSCCTTTKTYKTYKTYKTYKTYKTYKTYKTYKTYKTYKTYKTYKTYK 406  
 QY 1037 AGCCTGCGTGTGTTAGATCTTCTTCAAAATGTTGAAGCAGGAGTAAAGTCTTAATCA 1096  
 DB 407 TAAACTTCTGKSKTKMYMTTCTCTWTWYGTATTGAGRRGTARRAGACTAAANSA 466  
 QY 1097 GTGATTACTGCAAGAA 1112  
 DB 467 MTCTTCTCTTTGAA 482  
 RESULT 10  
 US-09-257-179-19  
 : Sequence 19, Application US/09257179  
 : Patent No. 6410709  
 : GENERAL INFORMATION:  
 : APPLICANT: Ruben et al.  
 : TITLE OF INVENTION: 29 Human Secreted Proteins  
 : FILE REFERENCE: P2015PI  
 : CURRENT APPLICATION NUMBER: US/09/257,179  
 : CURRENT FILING DATE: 1999-02-25  
 : EARLIER APPLICATION NUMBER: PCT/US98/17709  
 : EARLIER FILING DATE: 1998-08-27  
 : EARLIER APPLICATION NUMBER: 60/056,270  
 : EARLIER FILING DATE: 1997-08-29  
 : EARLIER APPLICATION NUMBER: 60/056,271  
 : EARLIER FILING DATE: 1997-08-29  
 : EARLIER APPLICATION NUMBER: 60/056,247  
 : EARLIER FILING DATE: 1997-08-29  
 : EARLIER APPLICATION NUMBER: 60/056,073  
 : EARLIER FILING DATE: 1997-08-29  
 : NUMBER OF SEQ ID NOS: 128



```
FILE REFERENCE: L0560/7005/ERP
CURRENT APPLICATION NUMBER: US/09/293,238B
CURRENT FILING DATE: 1999-04-16
PRIOR APPLICATION NUMBER: US 08/879,338
PRIOR FILING DATE: 1997-06-20
PRIOR APPLICATION NUMBER: US 08/663,731
PRIOR FILING DATE: 1996-06-14
PRIOR APPLICATION NUMBER: US 08/199,776
PRIOR FILING DATE: 1994-02-18
NUMBER OF SEQ ID NOS: 31
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 3927
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: CDS
LOCATION: (126)...(3665)
NAME/KEY: sig_peptide
LOCATION: (126)...(179)
NAME/KEY: mat_peptide
LOCATION: (180)...(3662)
US-09-293-238B-1

Query Match 2.68; Score 41.2; DB 4; Length 3927;
Best Local Similarity 67.4%; Pred. No. 0.26; 28; Indels 0; Gaps 0;
Matches 58; Conservative 0; Mismatches 0;

QY 1517 TGCATTGCGACACTGAAGATGATTAATCATATTTTATCTTTTAAATTAATCATGATT 1576
DB 3820 TGCATTGCGAAGTAACTTAGGAACAATTTGGTATTAAATAAATTACACTTTCT 3879

QY 1577 ATGCAAAAAAAAAAAAAAAAAAAAA 1602
DB 3880 TTGCAAAAAAAAAAAAAAAAAAAAA 3905

RESULT 14
US-08-199-776-1
Sequence 1, Application US/08199776
Patent No. 5594120
GENERAL INFORMATION:
APPLICANT: Brenner, Michael B.
APPLICANT: Parker, Christina M.
TITLE OF INVENTION: No. 5594120el integrin alpha subunit
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield and Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/199,776
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: B0801/7020
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3933 base pairs
TYPE: nucleic acid
TOPOLOGY: linear
```

```
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
CELL TYPE: mucosal lymphocyte
FEATURE:
NAME/KEY: CDS
LOCATION: 126..3662
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 180..3659
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 126..179
US-08-199-776-1

Query Match 2.68; Score 41.2; DB 1; Length 3933;
Best Local Similarity 67.4%; Pred. No. 0.26; 28; Indels 0; Gaps 0;
Matches 58; Conservative 0; Mismatches 0;

QY 1517 TGCATTGCGACACTGAAGATGATTAATCATATTTTATCTTTTAAATTAATCATGATT 1576
DB 3817 TGCATTGCGAAGTAACTTAGGAACAATTTGGTATTAAATAAATTACACTTTCT 3876

QY 1577 ATGCAAAAAAAAAAAAAAAAAAAAA 1602
DB 3877 TTGCAAAAAAAAAAAAAAAAAAAAA 3902

RESULT 15
US-08-663-731-1
Sequence 1, Application US/08663731
Patent No. 6057423
GENERAL INFORMATION:
APPLICANT: Brenner, Michael B.
APPLICANT: Parker, Christina M.
TITLE OF INVENTION: No. 6057423el integrin alpha subunit
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield and Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/663,731
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/199,776
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: B0801/7020
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3933 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
```

```

; MOLECULE TYPE: CDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; CELL TYPE: mucosal lymphocyte
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 126..3662
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 180..3659
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 126..179
; US-08-663-731-1

```

```

Query Match      2.6t; Score 41.2; DB 3; Length 3933;
Best Local Similarity 67.4t; Pred. No. 0.26;
Matches 58; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 1517 TGCATTTCGACACTGTAAAGATGATTATCATATATTTTATCTTTTAATTAAATCATGGATTT 1576
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3817 TGCATTGTGCAAAAAGTAAACTTAGGAAACATTGGTATTATTAATTACACTTTTCT 3876

Qy 1577 ATGCAAAAAAAAAAAAAAAAAAAAAA 1602
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3877 TTGCAAAAAAAAAAAAAAAAAAAAAA 3902

```

Search completed: August 2, 2004, 08:20:00  
Job time : 130 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2004, 05:04:23 ; Search time 1100 Seconds

(Without alignments)  
7140.755 Million cell updates/sec

Title: US-10-031-331b-39

Perfect score: 1602

Sequence: 1 cacacgctgatttaaccat.....aaaaaaaaaaaaaaaaaaaaa 1602

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 3222919 seqs, 2451570024 residues

Total number of hits satisfying chosen parameters: 6445838

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.\*

1: /cgn2\_6/ptodata/1/pubna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubna/US08\_PUBCOMB.seq.\*  
3: /cgn2\_6/ptodata/1/pubna/US09\_PUBCOMB.seq.\*  
4: /cgn2\_6/ptodata/1/pubna/US10\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubna/US11\_PUBCOMB.seq.\*  
6: /cgn2\_6/ptodata/1/pubna/US12\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubna/US13\_PUBCOMB.seq.\*  
8: /cgn2\_6/ptodata/1/pubna/US14\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubna/US15\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubna/US16\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubna/US17\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubna/US18\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/1/pubna/US19\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubna/US20\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubna/US21\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubna/US22\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubna/US23\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/1/pubna/US24\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/1/pubna/US25\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	834	52.1	1884	13	US-10-425-114-14370 Sequence 14370, A
2	822	51.3	1890	13	US-10-425-114-14613 Sequence 14613, A
3	743.4	46.4	1897	13	US-10-425-114-24087 Sequence 24087, A
4	726.4	45.3	2543	13	US-10-424-599-29759 Sequence 29759, A
5	688.8	43.0	1911	17	US-10-437-963-53228 Sequence 53228, A
6	676.2	42.2	1922	15	US-10-149-759-81 Sequence 81, Appl
7	505.2	31.5	1201	13	US-10-425-114-36489 Sequence 36489, A
8	446.4	27.9	1267	13	US-10-425-114-2412 Sequence 2412, A
9	397.2	24.8	825	13	US-10-425-114-25382 Sequence 25382, A
10	388.8	24.3	906	13	US-10-424-599-114653 Sequence 114653, A
11	358.2	22.2	594	17	US-10-021-323-14226 Sequence 14226, A
12	353.2	22.0	601	17	US-10-021-323-8285 Sequence 8285, A
13	329.4	20.6	525	17	US-10-021-323-14357 Sequence 14357, A
14	326.6	20.4	694	17	US-10-021-323-17175 Sequence 17175, A

15	316.4	19.8	586	17	US-10-021-323-13961 Sequence 13961, A
16	303.6	19.0	487	10	US-09-770-961-776 Sequence 776, Appl
17	290.6	18.1	500	17	US-10-021-323-17176 Sequence 17176, A
18	278.8	17.4	467	12	US-09-732-627A-3204 Sequence 3204, A
19	255.8	16.0	619	13	US-10-149-759-49 Sequence 49, Appl
20	238	14.9	557	13	US-10-424-599-97831 Sequence 97831, A
21	220.4	13.8	526	13	US-10-424-599-81253 Sequence 81253, A
22	209.8	13.1	1087	17	US-10-437-963-7552 Sequence 7552, Appl
23	194.4	12.1	494	10	US-09-770-961-584 Sequence 584, Appl
24	188	11.7	539	17	US-10-021-323-17177 Sequence 17177, A
25	184	11.5	591	17	US-10-021-323-3964 Sequence 3964, Appl
26	177.4	10.8	272	9	US-09-294-093B-1679 Sequence 1679, A
27	172	10.7	685	17	US-10-437-963-7551 Sequence 7551, A
28	169	10.2	290	9	US-09-294-093B-5022 Sequence 5022, A
29	164.2	10.2	477	17	US-10-021-323-6613 Sequence 6613, A
30	160.2	10.0	293	9	US-09-294-093B-160 Sequence 160, Appl
31	156.2	9.8	257	9	US-09-923-876-419 Sequence 419, Appl
32	152.6	9.5	257	11	US-09-923-876-419 Sequence 419, Appl
33	152.6	9.5	1416	17	US-10-602-268-18 Sequence 18, Appl
34	152.6	9.5	1534	17	US-10-602-268-18 Sequence 18, Appl
35	152.2	9.5	590	17	US-10-437-963-53196 Sequence 53196, A
36	150.8	9.4	508	13	US-10-424-599-73778 Sequence 73778, A
37	150	9.4	460	13	US-10-424-599-31677 Sequence 31677, A
38	144	9.0	261	9	US-09-923-876-984 Sequence 984, Appl
39	144	9.0	261	11	US-09-923-876-984 Sequence 984, Appl
40	142	8.9	290	9	US-09-294-093B-2860 Sequence 2860, A
41	140	8.7	299	9	US-09-294-093B-6039 Sequence 6039, A
42	139.6	8.7	5060	17	US-10-437-963-97616 Sequence 97616, A
43	134.8	8.4	274	9	US-09-923-876-3780 Sequence 3780, A
44	134.8	8.4	274	11	US-09-923-876-3780 Sequence 3780, A
45	132.6	8.3	314	9	US-09-294-093B-4550 Sequence 4550, A

## ALIGNMENTS

## RESULT 1

US-10-425-114-14370  
; Sequence 14370, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 14370  
; LENGTH: 1884  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB22-074-F3\_FLI  
US-10-425-114-14370

Query Match 52.1%; Score 834; DB 13; Length 1884;  
Best Local Similarity 74.1%; Pred. No. 2.68-219;  
Matches 1056; Conservative 0; Mismatches 370; Indels 0; Gaps 0;

Qy	1	CACACGCTGATTAAACCAATGAGCTATGATGCTCGAATCTCAAGCTTCGATCTTGAC	60
Db	256	CATTCCGCTGATCTGACTGTTGAAGCTATGATGCTGAGAGCTTCGATCTGAC	315
Qy	61	TAAGAGAGAGCTGAGATCTTTCATGCTTCGGCTCTGAAGAAATGCTCTTG	120
Db	316	AGGAAGAGAGCTGAGATCTTCTTTGCTCCCTCCATATGAAGGCAATCAGTGTG	375



421 TCTTTGTTCCATCAATCTGGGGATCACAAACGCAAGCAATCCACCACCTACCGTGA 480  
Db TCTTTGTTCCATCAATCTGGGGATCACAAACGCAAGCAATCCACCACCTACCGTGA 727  
481 CTTAGGCTCTCACTTAAGGCTCTCAAGAGTGTCTTTGCAAGATGATCTGGAAGCT 540  
Db CTTAGGCTCTCACTTAAGGCTCTCAAGAGTGTCTTTGCAAGATGATCTGGAAGCT 787  
541 TATGAGCTCTCCTACTAGTGCATATGATGATGATGATGATGATGATGATGATGAT 600  
Db TATGAGCTCTCCTACTAGTGCATATGATGATGATGATGATGATGATGATGATGAT 788  
601 CAGAACCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660  
Db CAGAACCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 847  
661 CAGAACCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 907  
662 CAGAACCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 967  
721 GGCCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
Db GGCCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1027  
781 GATTTGAAGCTCTGAGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 840  
Db GATTTGAAGCTCTGAGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1087  
841 TACATGCGGAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
Db TACATGCGGAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1147  
901 TCTTTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
Db TCTTTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1207  
961 TGCACCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
Db TGCACCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1267  
1021 CTGCATATGAGGCAAGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 1080  
Db CTGCATATGAGGCAAGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 1327  
1081 GGTAAAGTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
Db GGTAAAGTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1387  
1141 GCTTACATGAGGCAAGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 1200  
Db GCTTACATGAGGCAAGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 1447  
1201 AAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260  
Db AAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1507  
1261 CTACGAGGCAAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
Db CTACGAGGCAAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1567  
1321 GAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380  
Db GAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1627  
1568 AAGAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1426  
1381 GGTGAGGCAAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1673  
1628 GATGAGGCAAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT

## RESULT 3

US-10-425-114-24087

; Sequence 24087, Application US/10425114

; Publication No. US20040034808A1

GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10425,114  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 24087  
; LENGTH: 1897  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3607-019-D6\_FLI  
US-10-425-114-24087

Query Match 46.4%; Score 743.4; DB 13; Length 1897;  
Best Local Similarity 70.7%; Pred. No. 3e-194;  
Matches 1004; Conservative 0; Mismatches 416; Indels 1; Gaps 1;  
QY 1 CACACCGTTGATTTAAACCATTTGAAGCTATGATGCTCGATTTCTCAAGCTTTCTGATCTTGAC 60  
Db CACTCCGGGAGCTCAACCTCGAGGCCATTTATGCTCGACTCCCGTCCCGCGACTCGAC 299  
QY 61 AAAGAGACGCTCTGAGATCTTTCAATGCTTCCGCTCTTGAAGGAAATGCTCTTG 120  
Db AAGAGAGGCGCCCGAGGTTCTGCTTTCTCTTCTATATGAAGGAAATCTATACTG 359  
QY 121 GAATCTGGGCTGATTTGCTCTTTTACTGCTGAATTTGGCTGAGAAAGCTTGGCCAGTT 180  
Db GAGCTGGGAGCTGGAATAGCCGCTTTACTGCTGAATTTGGCTGAGAAATCTTGGCATGTT 419  
QY 181 ATTCTCTGATTTCAATTTGAGAGTGTATCAAGAGAGATGAATTAATCAATGAGCACTAC 240  
Db TTTCAGTGGATTTCTGAGAGTGTATTAAGAGATGAATTAATTAATGATCACTAT 479  
QY 241 AAAATGCTCAATTTATGCTGATGATGATTTCTCCACTCTCCACTCTCCACTCCACT 300  
Db GGCACACATCTTTATGCTGATGATGATTTATCCCGGACCTGATGATGATGATGATGAT 539  
QY 301 TCATGAGTGTATTTCTCAATTTGGTTACTCATGATTTCTCTGATGATGATGATGATGAT 360  
Db TCCATTCATCTGATTTTCAAACTGTTGCTGATGATTTCTTCAGATGATGATGATGATGAT 599  
QY 361 AATTGCTGCAAGAGATTTGAATGCTTGAAGCCAGGGGTTACATTTCTTCAGAGAA 420  
Db AAGTTGCTGCAAGAGATTTGAATGCTTGAAGCCAGGGGTTACATTTCTTCAGAGAA 659  
QY 421 TCTTGTTCATCAATCTCGGATCAAAACGCAAAAGCAATCCACCCTACCTACCTGAA 480  
Db TCTTGTTCATCAATCTCGGATCAAAACGCAAAAGCAATTTATCCCAACACTATGAGAA 719  
QY 481 CTTAGGTTCTTACACTAAGGCTTTCAAGAGTGTCTATTGCAAGATGATGATGATGATGAT 540  
Db CCAAGTTTATACCAAGGTTTAAAGATGCTCAAACTTTAATCAGATGATGATGATGATGAT 779  
QY 541 TATGAGCTCTCCTACTAGTGCATATGATGATGATGATGATGATGATGATGATGAT 600  
Db TCAAACTTTCTTTGATTAATCAATTTCAATGCTGATGATGATGATGATGATGATGATGAT 839  
QY 601 CAGAACCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660  
Db CAAACCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 899  
QY 661 CAGAACCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
Db AGTTTTTGGACAAATGTCAGTACAAAGCCCACTGGAATACCTACGCTATGAACTGATCTTT 959



Db 1613 ACACCATGCTGACGCTCAAGATATAACCAACATATTTTACAGATCAATTTTACAAGTGGTTGA 1672  
Qy 1073 AGCAGGAGGTAAGATTTCTAATCAGTGAATTTACTGCAAGAAAGCTGGTCCACCTCCACCTG 1132  
Db 1673 AGCTGGAGGAAATTTCTAATCACAGATTTACTGCAAAAGTCTGGAAGTCAATCTTTAG 1732  
Qy 1133 AATTCGCGCTTACATTTAAGCAGAGGGGATATGATCTCATGATGTAAGGAATATGGG 1192  
Db 1733 AATTTGCTGAGTACATATAAAGAAAGGAGGATATATCTCATGATTAAGCGTATAGG 1792  
Qy 1193 AGATGCTTAAAGATGCTGGATTTTGTGATGTTCTTGCCGAGATAGAACTGAGCAGTTCA 1252  
Db 1793 AGATGCTTGAAGATGCGGATTTGATGATGCTCATTTGCCGAGATCGAACTGATCAGTTG 1852  
Qy 1253 TTGCGATTTCTACGAAGGAACTAGAGCTGTTGAGAGGAAAGGATGTTGTCATTAGTG 1312  
Db 1853 TGAACAGCTACAGCAGAGTTAATGCTTGGAGAACAGAGGACGATTTTATGGTG 1912  
Qy 1313 ATTTCTTGAGGAGGATTAATGATGATTTGAGGATTTGGAATGATTAAGTTGCGGAGGA 1372  
Db 1913 ACTTCAGCGAGGAAAGCTACATGAGATTTGTCGAAAGATGGAAGCAAGCAGACCGGG 1972  
Qy 1373 CTGCAAGGGTGACCAACGATGGGCTGTTGCTTCCCAAGAAAGTGA 1422  
Db 1973 GTGCACTAGAGACGATGGGGCTTGTTCATTGCCAAGAAATTTGA 2022

## RESULT 5

US-10-437-963-53228  
; Sequence 53228, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yibua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boudharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 53228  
; LENGTH: 1911  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_5544C.1  
US-10-437-963-53228

Query Match 43.08; Score 688.8; DB 17; Length 1911;  
Best Local Similarity 69.48; Pred. No. 3.7e-179;  
Matches 967; Conservative 0; Mismatches 397; Indels 30; Gaps 1;  
Qy 79 ATCTTTCAATGCTTCGCTCTTGAAGGAAATGCTCTTGAATCTTGGGCTGGTAT 138  
Db 223 ATATTTGCTTTACTTCTCTCTTACGAGGAAATCAGTACTGGAACCTTGGTCTGGAATA 282  
Qy 139 GTCCTTTTACTGCTGAATTTGGCTGAGAAAGCTGGCAGGTTATTTGCTCTGATTTCAAT 198  
Db 283 GGTGCTTCACTGAGAACTAGTGAACAGCTGGGATGTTCTTGAATGATTTCAAT 342  
Qy 139 GAGAGTCTTCAAGAGAGATGAGTAAATCAATGGGCACTACAAAATGTCAGATTTATG 258  
Db 343 GAAAGTGTGATTAAGAGATGAAGCATTAAGCGGTCAACAGAGATGCTCTTATG 402  
Qy 259 TGTGCTGATGATCTTCTCCACTCTCAGTTTCCACCACTTCATTTGATGATGATATTC 318  
Db 403 TGTGGGATGTCACATGTCAGACCTGATGATGAGGATTAATCTCCATTTGATCTGATTT 462

Qy 319 TCCAAATGGTTACTCATGTATCTTTCTGATGAAGAGGTGGAATAATTTGGTTGAAGAATG 378  
Db 463 TCAAACTGGTTACTGATGATCTTTCTGAGCAGAGAGTTGAGAAGCTAGTAAAGAGAATG 522  
Qy 379 TGAATAATGGTTGAAGCCAGGGGGTTACATTTTCTTCAAGAGAACTCTTGTTCATCAATCT 438  
Db 523 GTAAGATGGCTAAAGGTTGGCGCTATATCTTCTTAGGGAATCTTGTTCATCAATCT 582  
Qy 439 GGGGATCAAAACCCAAAGCAATCCACCACTACCGTGAACCTAGTTCCTACACTAAG 498  
Db 583 GGAGATTCAAAAGGAAAGTGAATCTTACATTAACCGGAGCCAAAGTTTACACTAAG 642  
Qy 499 GCCTTCAAGAGTGTCTATTTGGAAGATGGAATCTGGAACCTCTTATGAGCTCTCCCTACT 558  
Db 643 GTGTTTAAAGAGTGTCAAGCTCTTGTATCAAGATGGGAATTCCTTTGAATCTCTGTACT 702  
Qy 559 AGCTGCAATATGTTGGAGCTTATGTGAGAAACAAAGAAACCAAGAACAGATAGTTGG 618  
Db 703 ACTTGAAGTGTGTTGGAGCTTACGTGAAAGCAAGAAATCAAAACCAAGATAGTTGG 762  
Qy 619 TTGTGGCAAAAGCTTGATTTCTAAGGATGATAAGGGTTCCAGCGATTTCTGACTAGC 678  
Db 763 CTATGGCAAAAGGTTGATTCACAGAGATCGGGGTTTCAAGATTTTGGCAATGAG 822  
Qy 679 CAGTACAAAGTGTAAATAGCATTTGCGATATGAGCGTATTTGGCCCTGGTTATGTTAGC 738  
Db 823 CAGTACAAAGCCAGTGGATATTAACGTATGAACGCATCTTTGGAGAAAGGCTTTGTGAGC 882  
Qy 739 ACTGGAGGATAT -----GAAACCAACCAAGAGATTT 768  
Db 883 ACTGTGGAATTTGATGCTTATTTTATCTTCGTTCAACCAAGAACTTACAAAGATTT 942  
Qy 769 GTGTCAATGCTGCACTTGAAGCTTGGCCGAGAGGCTCTGGATGTTGGTTGGAATTTGGT 828  
Db 943 GTGGACAGGCTGGATCTCAAACTGGCCAGAAAGCTTCTGTGATTTGGATGGAATTTGG 1002  
Qy 829 GGAGTGAATTTTACATGCGGAGACCTTTGATGTTGAGGTTGTTGGATTTGATCTCTCC 888  
Db 1003 GGCGTGAATTTATATGCTGACAAAGTATGATGTTCTATGTTGTTGTTGATTTGATCTCT 1062  
Qy 889 GTTAATGATTTCTTCTTGGCTTGAAGCTCTTATTTGGCTTAAATGCTGTTGAGTTT 948  
Db 1063 ATAAACATGTTTCTTGTGCACTGAGCGTCTATTTGGCGTAAAGTCTCAGTTGAGTT 1122  
Qy 949 GAGTGAAGATTTGACCAAGATAAACTACCTGATAACTCTTTTGAATGTCATCTATAGC 1008  
Db 1123 GAAGTCTGATTTGACCAAAAGACATACCCAGACACACGTTTGAAGTCTATCTACAGT 1182  
Qy 1009 CGTCAACCATTTCTGATATTCAGGACAAAGCTTGGTGTGTTAGATCTCTTCAAAAATG 1068  
Db 1183 CGTGAATCTCTTCAATCAAGATTAACCTCACTATTTTAAAGTTTCTTCAAGTGG 1242  
Qy 1069 TTGAAGCAGAGGTAAGTTTCTAATCAGTGAATTTCTGCAAGAAAGCTGGTCCACCTCA 1128  
Db 1243 CTCAAACTGGGGTAAAGTCTTAAATAGTGAATTTAGTGAATGCTTGGGAAACCTTCA 1302  
Qy 1129 CTTGAATTTCCCGCTTACATTAAGCAGAGGGAATGATCTCCATGATGTAAGAAATAT 1188  
Db 1303 GAAGATTTCCAGCTTACATTAAGCAAGGGTTATGACCTTTCACAGCTCAGGGCTTAC 1362  
Qy 1189 GGGCAGATGCTTAAAGATGCTGATTTGTTGATGTTCTTGGCGAGGATGAACTGAGCAG 1248  
Db 1363 GAACGATGCTTGAAGATGCTGTTTCCATGATGATCTTGTGAAGACCGCAGGATCAG 1422  
Qy 1249 TTCAATTCAGATTTCTAAGAGGAACTAGAGACTGTTGAGAAAGAAAGAGATGTTCTCAT 1308  
Db 1423 TTCTCGATGTTCTAGAGAGGAGCTTGTCTAAAGTTGAAAGACAAAGATGTTCTGTC 1482  
Qy 1309 AGTGATTTCTCAGAGGATTTACAAATGATGTTGAGGTTGGAATGATAGTTGGG 1368  
Db 1483 TCTGATTTTCAAGCAGAGGATCTAGAGCCTATGATGATGAGGCAAAACTTCAA 1542

QY 1369 AGGACTGCCAAGGCTGAGCAAGATGGGGTCTGTTGGTTCGACAGAGAGTGAAGATC 1428  
DB 1543 AGGAGTTCTGCTGGTGGCAGAGGTGGGGCTGTTTCATCGCGACCAAGTGAATCAAG 1602  
QY 1429 AGTTGGCGGCACTGG 1442  
DB 1603 ATGCGCAGCTGGG 1616

## RESULT 6

US-10-149-759-81  
; Sequence 81, Application US/10149759  
; Publication No. US20030157592A1  
; GENERAL INFORMATION:  
; APPLICANT: Lerchl, Jens  
; APPLICANT: Renz, Andreas  
; APPLICANT: Reinhardt, Thomas  
; APPLICANT: Reindl, Andreas  
; APPLICANT: Cirpus, Petra  
; TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins  
; TITLE OF INVENTION: involved in the synthesis of tocopherols and  
; TITLE OF INVENTION: carotenoids.  
; FILE REFERENCE: BASF/NAE 1333/99 PCT/US  
; CURRENT APPLICATION NUMBER: US/10/149, 759  
; PRIORITY FILING DATE: 2002-10-17  
; PRIOR APPLICATION NUMBER: PCT/EP/00/12698  
; PRIOR FILING DATE: 2000-12-14  
; NUMBER OF SEQ ID NOS: 82  
; SOFTWARE: WordPerfect version 6.1  
; SEQ ID NO 81  
; LENGTH: 1962  
; TYPE: DNA  
; ORGANISM: Physcomitrella patens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (367)..(1842)  
; OTHER INFORMATION: 78\_ppprotol\_092\_e12-260rev  
US-10-149-759-81

Query Match 42.2%; Score 676.2; DB 15; Length 1962;  
Best Local Similarity 68.1%; Pred. No. 1.2e-175;  
Matches 974; Conservative 0; Mismatches 448; Indels 9; Gaps 2;  
QY 1 CACACGTTGATTAACCATTAAGCTATGATGCTGATCTCAAGCTCTCTGATCTTGAC 60  
DB 415 CATTCTGTGGAGCTAGCGTTGAGCAATGATGCTTGTGAGGCTCCAACTCGAT 474  
QY 61 AAAGAAGAAAGTCTGAGATCTCTTCAATGCTTCGGCTCTTGAAGGAATGCTCTTG 120  
DB 475 AAAGAAGAAAGCAACCGAGATTTGCTGCTGTGCGCCCATATGAAACAGGATGTCATG 534  
QY 121 GAATCTGGGGCTGATGTTGGTGTGTTTACTGGTGAATGCTGAGAAAGCTGGCCAGGT 180  
DB 535 GAGCTCGAGCAGGCACTGGTGGTGTGCTTACTGGTGAAGCATGCAAGCTCATGTG 594  
QY 181 ATGCTCTGATTTCAATGAGTGTCTATCAAGAGAGTGAATCAATGGCACTAC 240  
DB 595 CTGCTGATGATTTCAATGAGATCTCTATCAAGAGAGTGAATGAGTGAAGTCACTAC 654  
QY 241 AAAAATGCAAGTTTATGTTGCTGATGCTCTCTCCACTCTCAAGTTCCCAACACAT 300  
DB 655 AACACATGCAATTTCAATGTCGATGTCAGCTCTCCAGACCTGAAATATTGCGAGGT 714  
QY 301 TCAATGAGTGAATTTCTCAATTTGGTGTCTCATGATCTCTTCTGATGAAGAGGTGGA 360  
DB 715 TCTGCGATCTGGTGTGTTCAATTTGGCTTCTCATGTTCTGCTGACGAGAGGTAAA 774  
QY 361 AATTGGTTGAAGATGTTGAATGCTTGAAGCAGGCTTCAATTTCTTCAGGAA 420  
DB 775 GGCCTAGATCAGCGTGTGAGGTGCTCAGGCTGAGATGATCTTCTCAGAGAA 834  
QY 421 TCTTGTTCATCAATCTGGGATCACAAGCGAAAGCAATCCCAACCACTACCGTGA 480

DB 835 TCCTGCTTCCACAGTCAAGAGATCAAGCGAAAGAAATCTCTACTACTACCTCGTAA 894  
QY 481 CCTAGGTTCTACACTAAGCGCTTCAAGAGTGTCTATTGCAAGATGATCTGGAATCTCT 540  
DB 895 CCAACAGGTACACGAACATCTTCCAGCAGGCTTACATCGAAGAGG---ATGGGCTCTAT 951  
QY 541 TATGAGCTCTCCCTACTTATAGCTGCAAAATGATTTGAGCTTATGTCAGAAACACAAAC 600  
DB 952 TTCAAGTTTGAATGGTTCGATGCAAAATGTTGCGCATATACGTGCGCAATTAAGAAAT 1011  
QY 601 CAGAACAGATTTAGTTGGTGGCAAAAGTTGATTTCTAA-----GGATGATAAGGG 654  
DB 1012 CAAACCCAGGTGTGTTGGTTATGAGAGAAAGTTTCACTCGGATGGACCTGAGAGCGTGT 1071  
QY 655 TTCCAGCGATTTCTGGATACTAGCAGTACAAAGTGTAAATAGCATTTCTCCGATATAGCGT 714  
DB 1072 TTCAGAAAGTTTTCGACACCCCAACAGTACACGTCACTCGAATCTCTCGTTACAGCGT 1131  
QY 715 GTATTGGCCCTGTTATGTTAGCCTGGAGGATGAAACCAACCAAGAGTTTGTGTCA 774  
DB 1132 ATTTTGGAGAGGATTTGTAGCAGGCTGGAATCGAAACCAACCAAGAGTTTGTAGT 1191  
QY 775 ATGCTGGAATTGAAGCTCGCCAGAAAGTCTCTGGATTTGGTGTGGAATTTGGTGGAGT 834  
DB 1192 ATGCTGGAATTGAAGCTCGCCAGGACAGCGTCTCTTGAAGTTGGAATGCGGATCGAGTGGT 1251  
QY 835 GACTTTTACATGGCGAGACCTTTGATGTTGAGGTTTGGATTTGATCTCTCCGTTAAT 894  
DB 1252 GATTTCTACATGGCGAGAAATATGATGCTGAAGTTTGGATGCACTCTCTTAAT 1311  
QY 895 ATGATTTCTTGGCTTGGAGCTTCTAATGGCTTAAATGGCTTAAATGGTGTGAGTTGAGTA 954  
DB 1312 ATGATTTCTTGGCTTGGAGCTTCTAATGGCTTAAATGGCTTAAATGGTGTGAGTA 954  
QY 955 GCAGATTTGCAACCAAGATAAATCTACCTGATAATCTCTTTGATGTCATCTATAGCGTGCAC 1014  
DB 1372 GGGGATTTGCAACCAAGATAAATCTACCTGATAATCTCTTTGATGTCATCTATAGCGTGCAC 1431  
QY 1015 ACCATTTCTGATATTTAGGCAAGCTCGGTTGTTAGATCTCTTACAAATGGTTGAG 1074  
DB 1432 ACCATTTCTGATATTTAGGCAAGCTCGGTTGTTAGATCTCTTACAAATGGTTGAG 1491  
QY 1075 CAGGAGGTAAAGTTCTAATCAGTGAATTTACTGCAAGAAAGCTGGTCCACCTCACTGAA 1134  
DB 1492 CTTGAGGTCTGGTGTCTTATCAGTGAATTTACTGTAAGCTCCACAACTCCGTCGCGGAG 1551  
QY 1135 TTGCGCGCTTACATTTAAGCAGAGGAGATATGATCTCCATGATGTAAGAAATATGGCAG 1194  
DB 1552 TTGCTGATACATTTAGCAGAGGAGTATGATCTCTATAGCTTCAAGATTAAGAGAG 1611  
QY 1195 ATGCTTAAAGATGCTGAATTTGTTGATGTTCTTGGCGAGGATAGAACTGAGCAGTTCAAT 1254  
DB 1612 ATGCTGGAAGATGCGGTTTGTGGAAGTGTGCGAGAGGACCGCAGGATCAGTTCAAT 1671  
QY 1255 CAGATTTCTACGGAAGAACTAGAGACTGTTGAGAGGAAAGAGATGTTCTATTAAGTGAAT 1314  
DB 1672 GAAGTTTACAGAGGAGCTAGCCACCACTGAAGCAGGTCTGTCAGCAGTTCTCAACAGAT 1731  
QY 1315 TTCTGAGGAGGATTAACATGATTTGAGAGTTGGAAGTTGGAATGATTAAGTTGCGGAGCT 1374  
DB 1732 TTCTGAGGAGGATTAACATGATTTGAGAGTTGGAAGTTGGAATGATTAAGTTGCGGAGCT 1791  
QY 1375 GCCAAGGAGTGAACCAAGTGGGCTCTGTTGTTGTCGCAAGAGAGAGTGAAGA 1425  
DB 1792 TCGAATGCAACCAAGAGTGGGAGCTCTCTCATAGCTTACAGGCAATTATGA 1842

## RESULT 7

US-10-425-114-36489  
; Sequence 36489, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong









Db 61 GTGGCAAGCTGGGACATAAACCTGGCCAGAAAGTACTGGAGTGTGGTACTGGG 120  
Qy 829 CGAGGTGACTTTTACATGGCGGAGACCTTTGATGTTGGAGTGTGGATTTGATCTCTCC 888  
Db 121 GAGGTGACATTTACATGGCAGAAATTTGATGTTGGAGTGTGGATTTGATCTCTCC 180  
Qy 889 GTTAATATGATTTCTTGGCCCTTGGAGGCTTCTAATGGGCTTAATGTCGTGTAGATT 948  
Db 181 AATAACATATTTCTTGGCCATTTGAACGCTGCTAATGGACTCAATGCTGTGTAGATT 240  
Qy 949 GAGGTAGCAGATTGACCAAGATATAACTACCTGATATACTTTTATGATCTATATAGC 1008  
Db 241 GAGTGTGGCGATTGCACTTAAACCAATTTCTCTGTGATATACATTTGATGATCTATTC 300  
Qy 1009 CGTGACACATCTGCAATTTACGACAGAGCTGCTGTTTATAGTCTTCTACAAATGG 1068  
Db 301 CGTGACACATCTGCAATTTACGACAGAGCTGCTGTTTATAGTCTTCTACAAATGG 360  
Qy 1069 TTGAAGCCAGGAGGTAAAGTTCTAATCAGTGAATTAAGTCTGCAAGAAAGCTGCTCCACCTCA 1128  
Db 361 TTGAAGCCAGGAGGTAAAGTTCTAATCAGTGAATTAAGTCTGCAAGAAAGCTGCTCCACCTCA 420  
Qy 1129 CTTCAATTCGCGCTTACATTAAGCAGAGGAGTATGATCTCCATGATCTTAAGGATAT 1188  
Db 421 TTAGGATATCTGAGTATATAAGAAAGGAGGATATATATCTGATGATGAACATAT 480  
Qy 1189 GGGCAGATCTTAAGATCTGGATTTGTTGATGTTCTTCCGAGGATAGACTGAGGAG 1248  
Db 481 TGTGGATGCTCGAATGCTGGATTTGATGATGTTGTTGCGAGGATGCAATATTTG 540  
Qy 1249 TTTCTTCGAGTTCTACGGAAGGAACTAGAGACTGTTGAGAAAGAAAGATGTTTCAAT 1308  
Db 541 TTTCTTCGAGTTCTACGGAAGGAACTAGAGACTGTTGAGAAAGAAAGATGTTTCAAT 600  
Qy 1309 AGTGAATTTCTGAGGAGATCAATGACATGTTGAGGATGAGATGATGATTTGGG 1368  
Db 601 GATGACTTCTGGAGAGAGACTACATGATGATGATGATGATGATGATGATGATGAT 660  
Qy 1369 AGGACTGCCAGGCTGACAGATGAGGCTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1428  
Db 661 CGGGGGGCGAGATGTTGAAACAAATATGAGGCTTGTTCATGTTGTTGTTGTTGTTGTTGTT 720

## RESULT 11

US-10-021-323-14225  
; Sequence 14225, Application US/10021323  
; Publication No. US20040123340A1  
; GENERAL INFORMATION:  
; APPLICANT: Deikman, Jill  
; APPLICANT: Feng, Paul C.C.  
; APPLICANT: Fincher, Karen L.  
; APPLICANT: Ziegler, Todd E.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(52274)B  
; CURRENT APPLICATION NUMBER: US/10/021,323  
; CURRENT FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: US 60/255, 619  
; PRIOR FILING DATE: 2000-12-14  
; NUMBER OF SEQ ID NOS: 17880  
; SEQ ID NO 14225  
; LENGTH: 599  
; TYPE: DNA  
; ORGANISM: Gossypium hirsutum  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(599)  
; OTHER INFORMATION: unsure at all n locations  
US-10-021-323-14225

Query Match

22.2%; Score 355.2; DB 17; Length 599;

Best Local Similarity 75.7%; Pred. No. 2.9e-87;  
Matches 452; Conservative 0; Mismatches 144; Indels 1; Gaps 1;  
Qy 761 AAGATTTGTGTCAATGCTGGACTTGAAGCTCGCCAGAGGCTCTTGGATGTTGGTGTG 820  
Db 1 AAGATTTGTGTGCGCAATTTGATCTTAAGCTCTGCGCA-AAGGTTCTTATGATGTTGGCTGTG 59  
Qy 821 GAATTTGGTGGAGGTGACTTTTACATGCGCGAGACCTTTTGAATGTTGAGGTTGTTGATTTG 880  
Db 60 GCATAGTGGAGGTGACTTTTATATGCGCAAGAAATTTGAGTCTATGTTGGCATCG 119  
Qy 881 ATCTCTCGTTAATATGATTTCTTTCCTTGGCTTGGAGGTTCTATTTGGGCTTAAATGCTG 940  
Db 120 ATCTCTCGTTAATATGATTTCTTTCCTTGGCTTGGAGGCTTAAATGCTGCTG 179  
Qy 941 TTGAGTTTGGGTAGCAGATTGCAACCAAGATAAATACCTCTGATTAATCTTTTGTATGTC 1000  
Db 180 TGGAAATTTGAAGTTGCTGATTTGCACTACAAAGACTTATCCAGATAACAGCTTTTGTATGTTA 239  
Qy 1001 TCTATAGCGCTGACACCATTTCTGATATTCAGCAGAGCTCGGTTGTTAGATCTTCT 1060  
Db 240 TCTAGCGCTGACACCATTTCTGATATTCAGCAGAGCTCGGTTGTTAGATCTTCT 299  
Qy 1061 ACAAATGTTGAGCAGAGGAGGTTAAAGTTCTTATCAGTGAATTAAGTCTGCAAGAAAGCTGTC 1120  
Db 300 ACAAATGTTGAGCAGAGGAGGTTAAAGTTCTTATAGTGAATTAAGTCTGCAAGAAAGCTGTC 359  
Qy 1121 CACCTCTCACTCTGAAATTTCCCGCTTTACATTAAGCAGAGGAGGATATGATCTCCATGATGTA 1180  
Db 360 CTTCTATCGCCGAAATTTGCAAGATGATCAATCAAGCAGAGAGGCTATGATCTCCATGATGTA 419  
Qy 1181 AGGATATGGCAGATCTTAAAGATGCTGATTTGTTGATGTTCTTTCGCGAGGATGAA 1240  
Db 420 AAGCTTATGGCAGATCTTAAAGATGCTGATTTGTTGATGTTCTTTCGCGAGGATGAA 479  
Qy 1241 CTGACGATTCATTCGAGTTTCTACGAAAGAACTAGAGCTGTTGAGAAAGAAAGGATG 1300  
Db 480 CCGATCAGTTTCAATCAAGTTCTGAGCGTGAATTTAGCAAAAGTTGAGAAAGAAAGGATA 539  
Qy 1301 TGTTCATTTAGTGAATTTCTGAGGAGGATTTACAAATGATGATTTGAGGTTGGAATG 1357  
Db 540 CTTCTATCGCCGAAATTTCTCAGGAGAGCTTACGAGATGATGATTTGAGGATGGAATG 596

## RESULT 12

US-10-021-323-8285  
; Sequence 8285, Application US/10021323  
; Publication No. US20040123340A1  
; GENERAL INFORMATION:  
; APPLICANT: Deikman, Jill  
; APPLICANT: Feng, Paul C.C.  
; APPLICANT: Fincher, Karen L.  
; APPLICANT: Ziegler, Todd E.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(52274)B  
; CURRENT APPLICATION NUMBER: US/10/021,323  
; CURRENT FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: US 60/255, 619  
; PRIOR FILING DATE: 2000-12-14  
; NUMBER OF SEQ ID NOS: 17880  
; SEQ ID NO 8285  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Gossypium hirsutum  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3828-008-Q1-K6-B6  
US-10-021-323-8285

Query Match

22.0%; Score 353.2; DB 17; Length 601;

Best Local Similarity 75.2%; Pred. No. 1e-86;  
Matches 453; Conservative 0; Mismatches 148; Indels 1; Gaps 1;

QY 753 AACCAACAAAGAGTTTGTGTCAATGCTGGAATCTGAAGCTGCGCAGAGGTCCTGGATG 812  
DB 1 AAAACAACAAGAGTTTGTGGGAGTTAGATCTTAAGCTTGGCA- AAGTCTTAGATG 59  
QY 813 TGGTGTGGATTTGGTGGAGTGACCTTTACATGGCGGAGACCTTTGATGTTGAGGTTG 872  
DB 60 TGGCTGTGGCAATTTGGTGGAGTGACATTTATATGGCTGAGGAAATTTGATGTTGATGTT 119  
QY 873 TGGATTTGATCTCTCCGTTAATATGATTTCTTTTGGCCCTTGAGCGGTTCTATTGGGCTTAA 932  
DB 120 GGCGCAACGACCTCTCTGTAACATGATATCTTTGCTTTGACAGGCTACTGGACTGA 179  
QY 933 ATGCTGTGTGAGTTTGGAGTAGCAGATTGCACCAAGATAAATCTACCTTGATTAATCTTT 992  
DB 180 ATGCTGAGTGGAAATTTGAGTTGCTGATTCACCAAGAGGTTTATCCCGACACAGTTT 239  
QY 993 TGATGTCTATATAGCGGTGACACCATTTCTGCAATATTCAGGACAGGCTGCGTTGTTAG 1052  
DB 240 TGATGTATCTACAGCGGTGACATTTCTTACATATTCATGACAAACCTGCACTGTTAG 299  
QY 1053 ATCTCTTACAAATGTTGAGCCAGGAGGTAAGTTTCTAATCAGTGATTAAGTCAAGAA 1112  
DB 300 ATCTCTTACAAATGTTGAGCCAGGAGGCAAACTCTCTAATGATTAATCTGCAAAAG 359  
QY 1113 AGCTGGTCCGCTCACTGCAATTTGGCGCTTACATTTAGCAGAGGATATGATCTCA 1172  
DB 360 TTCCAGAGCTCCATCCAGGAGTTTGTGAGTATATCAGCAGAGGCTATGATCTTCA 419  
QY 1173 TGAATGAAGGAAATATGGCAGATGCTTTAAGATGCTGGATTTGTTGATGTTCTTGGCGA 1232  
DB 420 TGAATGAATATCATATGACAGATGCTTGAGGATGCTGGTTTGTGATGTTGATTTGCGA 479  
QY 1233 GATAGACTGAGGAGTTTCAATTCAGTTTCTACGAGGAACTAGACATGTTGAGAGGA 1292  
DB 480 GATGCAACGATGATCTTACAGTTCTCGCGGTGATTCGCAAGTGGAGGAAGA 539  
QY 1293 AAGAGTGTCTATGATGATTTCTGAGGAGGATTAATGATGATTTGAGGTTG 1352  
DB 540 GAAGGATGCAATTCATCTCTGACTTCTCTAAGGAAGACTATGATGAATAGTTGGTGGATG 599  
QY 1353 GA 1354  
DB 600 GA 601

## RESULT 13

US-10-021-323-14357  
; Sequence 14357, Application US/10021323  
; Publication No. US20040123340A1  
; GENERAL INFORMATION:  
; APPLICANT: Deikman, Jill  
; APPLICANT: Feng, Paul C.C.  
; APPLICANT: Fincher, Karen L.  
; APPLICANT: Ziegler, Todd E.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(52274)B  
; CURRENT APPLICATION NUMBER: US/10/021.323  
; PRIOR FILING DATE: 2001-12-12  
; PRIOR FILING DATE: 2000-12-14  
; NUMBER OF SEQ ID NOS: 17880  
; SEQ ID NO 14357  
; LENGTH: 525  
; TYPE: DNA  
; ORGANISM: Gossypium hirsutum

Query Match 20.6%; Score 329.4; DB 17; Length 525;  
Best Local Similarity 76.9%; Pred. No. 3.6e-80;  
Matches 402; Conservative 0; Mismatches 121; Indels 0; Gaps 0;  
US-10-021-323-14357

QY 750 TGAACCAACAAAGAGTTTGTGTCAATGCTGGAATCTGAAGCTGCGCAGAGGTCCTGGA 809  
DB 3 TGAACAACAAGAGTTTGTGGGAGTTAGATCTTAAGCTTGGCAAAAGGTCCTAGA 62  
QY 810 TGGTGTGGATTTGGTGGAGTGACCTTTACATGGCGGAGACCTTTGATGTTGAGGTT 869  
DB 63 TGGTGTGGATTTGGTGGAGTGACATTTATATGGCTGAGGAAATTTGATGTTGATGTT 122  
QY 870 TGGTGGATTTGATCTCTCCGTTAATATGATTTCTTTTGGCCCTTGAGCGGTTCTATTGGGCT 929  
DB 123 TGTGGCAACGACCTCTCTGTAACATGATATCTTTGCTTTGACAGGCTACTGGACT 182  
QY 930 TAAATGCTGCTGTCAGTTTGGAGTAGCAGATTGCACCAAGATAAATCTACCTTGATACTC 989  
DB 183 GAAATGCTCAGTGGAAATTTGAGTTGCTGATTCACCAAGAGGTTTATCCCGACACAG 242  
QY 990 TTTGATGTCTATATAGCGGTGACACCATTTCTGATATTCAGGACAGGCTGCTGTTT 1049  
DB 243 TTTTGTATGTTATCTACAGCGGTGACATTTCTTACATTTGATGACAAACCTGCACTGTT 302  
QY 1050 TAGATCTCTTACAAATGTTTGAAGCCAGGAGGTAAGTTTCTAATCAGTGATTAAGTCA 1109  
DB 303 TAGATCTCTTACAAATGTTTGAAGCCAGGAGGCAAACTCTCTAATGATGATTAAGTCA 362  
QY 1110 GAAAGCTGGTCCACCTCACTGAAATTTGGCGCTTACATTAAGCAGAGGAGGATATGATCT 1169  
DB 363 AAGTTCCAGACTCCATCCAGGAGTTTGTGAGTATATCAGCAGAGGAGGCTATGATCT 422  
QY 1170 CCATGATTAAGGAATATGGCAGATGCTTTAAGATGCTGGAATTTGTTGATGTTCTTGC 1229  
DB 423 TCATGATTAATATCATATGACAGATGCTTTGAGGATGCTGGTTTGTGATGTTGATTTTGC 482  
QY 1230 CCAGATAGAACTGAGCAGTTTCAATTCAGTTTCTACCGAGAGAA 1272  
DB 483 AAGAGATCGAACCGATCACTTCTTACAGTTCTCGCGGCTGAA 525

## RESULT 14

US-10-021-323-17175/c  
; Sequence 17175, Application US/10021323  
; Publication No. US20040123340A1  
; GENERAL INFORMATION:  
; APPLICANT: Deikman, Jill  
; APPLICANT: Feng, Paul C.C.  
; APPLICANT: Fincher, Karen L.  
; APPLICANT: Ziegler, Todd E.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(52274)B  
; CURRENT APPLICATION NUMBER: US/10/021.323  
; PRIOR FILING DATE: 2001-12-12  
; PRIOR FILING DATE: 2000-12-14  
; NUMBER OF SEQ ID NOS: 17880  
; SEQ ID NO 17175  
; LENGTH: 694  
; TYPE: DNA  
; ORGANISM: Gossypium hirsutum

Query Match 20.4%; Score 326.6; DB 17; Length 694;  
Best Local Similarity 74.1%; Pred. No. 2.6e-79;  
Matches 413; Conservative 0; Mismatches 144; Indels 0; Gaps 0;  
US-10-021-323-17175  
QY 866 AGTGTGTGATTTGATCTCTCCGTTAATATGATTTCTTTGCTGAGGCTTCTATTG 925  
DB 694 ATGTTGTGGCATTCATCTGTCACATATGAATTTTCTTCTGAAAGCCCATG 635  
QY 926 GCTTAAATGCTGTTGAGTTGAGTAGCAGATTGACAGATAAATACCTGATA 985

```
Db 634 CACTCGATGCTCTGTGGAAATTTGCACTTGTGATGCTACACAGACTTATCCAGATA 575
Qy 986 ACTCTTTTGTGATGCTATATAGCCGTCACACACCTTCTGCATATTCAGACAGCCCTGCGT 1045
Db 574 ACAGCTTTGATGATATCTATACAGCCGTCACACAGATTTTACACATTTCAAGACAGCCCTGCAC 515
Qy 1046 TCTTTAGATGCTCTTACAAATGTTGAAGCCAGAGGTAAAGTTCTTAATCAGTGAATTA 1105
Db 514 TATTTAGATATTTTACAAATGTTGAAGCCAGGCGGCAAAAGTTCTTATAAGTGAATTA 455
Qy 1106 GCAAGAAAGCTGCTCCACCTCACTCAATTTGCGCCGCTTACATTAAGCAGAGGCGATATG 1165
Db 454 GCAAAAGCCAGAGGCTTCTGCGCCGCAATTTGCAGAGTACATCAAGCAGAGAGGCTATG 395
Qy 1166 ATCTCCATGATGAAGAAATATGGCAGATGCTTAAGATGCTGATTTGTTGATGTTTC 1225
Db 394 ATCTCCATGATGAAGAAATATGGCAGATGCTTAAGATGCTGTTGTTGATGATGCTCA 335
Qy 1226 TTGCGAGGATAGAACTGAGCAGTTTCAATTCGAGTTCTACGGAAGGAACTAGAGACTGTTG 1285
Db 334 TTGCGAGGATAGAACTGAGCAGTTTCAATTCGAGTTCTACGGAAGGAACTAGAGACTGTTG 275
Qy 1286 AGAAGGAAAGGATGTTGTTTCAATTCGAGTTCTACGGAAGGAACTAGAGACTGTTG 1345
Db 274 AGAAGGAAAGGATGTTGTTTCAATTCGAGTTCTACGGAAGGAACTAGAGACTGTTG 215
Qy 1346 GAGTTTGGATGATTAAGTTGCGGAGGACTGCGAAGGTTGAGCAACGATGGGTTCTGTTGG 1405
Db 214 GGGGATGGAGCAAGGTAATAGGACCAAGATTTGGGGAGCAGAGATGGGGTCTTTTCA 155
Qy 1406 TTGCCAAGAAAGTGA 1422
Db 154 TTGCCAACAATAACTAA 138
```

## RESULT 15

```
US-10-021-323-13961
; Sequence 13961, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Peng, Paul C.C.
; APPLICANT: Pincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 13961
; LENGTH: 586
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(586)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: LIB3829-022-Q6-K6-H6
US-10-021-323-13961
```

```
Query Match 19.8%; Score 316.4; DB 17; Length 586;
Best Local Similarity 75.3%; Pred. No. 1:5e-76;
Matches 406; Conservative 0; Mismatches 132; Indels 1; Gaps 1;
Qy 337 TATCTTTCTGATGAAGAGGTGAAATTTGGTTGAAAGAAATGTTGAAATGTTGAAAGCCA 396
Db .. 45 TGTATCTCTCANTAAAGAGTTGAGAAATTTGGCGG-AGGATGCTCAAGTGTGAAGTT 103
Qy 397 GGGGGTTACATTTTCTTCAGAGAACTCTTTCATCAATCTGGGATCACAAGCGCAA 456
```

```
Db 104 GGTGACACATATTTTCTTCAGGAGTCTTGCTTTCATCAATCTGGTGNATTCAGAGAAAG 163
Qy 457 AGCAATCCCAACCACTACCGTGAACTAGGTTCTTACACTAAGGCCCTTCAAGAGTGTCAAT 516
Db 164 AACAAACCAACTCACTATCGCGAGCCAGATTTTACACTAAGGTCTTTAAAGAAATGCCAA 223
Qy 517 TTGCAAGATGGATCTGGAATCTCTTATGAGCTCTCCCTACTTCTAGCTGCAAAATGATTTGA 576
Db 224 GCAACTGATGATTTCTGGAATTCATTTGAATTTCTCTTGTGGCTGCAAGTGCATTTGA 283
Qy 577 GCTTATGTACAAACAAGAAACCAAGAACAGATAGTTGTTGTTGGCAGAAAGTTGAT 636
Db 284 GCTTATGTTAAGACAAAGAAAGATCAGATCAGATTTGCTGGATCTTCCAAAAGTTGTT 343
Qy 637 TCTAAGGATGATAAGGGGTTCCAGCGGATTTCTGGATCTAGCCAGTACAAAGTGTAAATAGC 696
Db 344 TCAGATAATGATAAGGGATTTCCAGCGCTTCTTGGATTTCTGTTTCACTACAAATCTAATAGC 403
Qy 697 ATTCTCGGATATGAGCGTGTATTTGGCCCTGTGTTATGTTAGCACTGGAGGATATGAACC 756
Db 404 ATTCTCGCTTATGAGCGTGTCTTTGGACCAAGGATATGTCAGCACAGGAGGATTCAGACA 463
Qy 757 ACCAAAGATTTGTGTCATGCTGGACTTGAAGCCCTGGCCAGAGGTCTCTGGATTTGGT 816
Db 464 ACAAGAAATTTGTGGGGAAGTTAGATCTTTAAGCCCTGGCCAAAGGTCTTAGATTTGGC 523
Qy 817 TGTGGAATTTGCTGGAGGTGACTTTTACATGGCGGAGACCTTTTCATGTTTGAGGTTGTTGG 875
Db 524 TGTGGCATTTGCTGGAGGTGACATTTATATGCTGAGGAAATTTGATGTTTCAATGTTTGGG 582
```

Search completed: August 2, 2004, 08:38:29  
Job time : 1102 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2004, 04:40:09 ; Search time 4442 Seconds

(without alignment)  
10769.746 Million cell updates/sec

Title: US-10-031-331B-39

Perfect score: 1602

Sequence: 1 cacacccgttgatttaaccat.....aaaaaaaaaaaaaaaaaaaa 1602

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:

1: em\_estba:\*

2: em\_estham:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_estc:\*

9: gb\_estl:\*

10: gb\_est2:\*

11: gb\_estc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	732.8	45.7	2333	11	AY103779	Zea mays
2	654	40.8	713	10	BE231445	BR231445 S650815 S
3	566.6	35.4	1012	11	AY105261	AY105261 Zea mays
4	538.6	33.6	1211	10	BE574969	BE574969 E07 Triph

5	465.6	29.1	1022	14	CK206655	FGAS01825
6	464.6	29.0	767	12	BM408047	EST582374
7	463.2	28.9	781	14	CF208434	CA280003
8	461.6	28.8	869	14	CF513471	CA280007
9	452	28.2	1115	14	CK166076	FGAS05012
10	451.4	28.2	753	14	CB970167	CB970167
11	440.6	27.5	720	14	CF508254	USDA-EPI
12	435.2	27.2	601	12	BM302134	MCA047F06
13	429	26.8	822	10	BE040284	OE03A05
14	427.2	26.7	706	13	BQ120291	EST605867
15	426	26.6	733	13	BQ965516	QHB22C02
16	421.8	26.3	731	13	BQ855322	QHB22C08
17	421.6	26.3	805	10	BE040460	OE06G12
18	416.4	26.0	721	14	CF482195	PO11_5_B0
19	415.6	25.9	941	14	CK265554	EST711632
20	413.2	25.8	682	13	BQ856500	QHB4H12.Y
21	410.4	25.6	703	14	CB972747	CAB30001
22	409	25.5	736	13	BQ852610	QHB18H22
23	408.6	25.5	750	12	BM406709	EST581036
24	407.8	25.5	718	14	CA174643	SCJFST101
25	407	25.4	527	13	BQ587932	E012339-O
26	404.6	25.3	700	14	CA289955	SCAGFL801
27	403	25.2	1127	10	BE420987	HMM004.F0
28	401	25.0	771	14	CD920161	G508_116A
29	400.2	25.0	771	14	CA327433	SCACSD201
30	399.2	24.9	848	14	CF513552	CABud0007
31	398.8	24.9	731	13	CA100170	SCVFCU604
32	398.6	24.9	806	14	CK195203	FGAS00364
33	394.6	24.6	689	13	BQ119788	EST605364
34	393	24.5	664	13	BQ871545	QGI12C12
35	391.4	24.4	802	12	BJ573651	BJ573651
36	390.2	24.4	745	13	BQ762263	EBR001.SQ
37	389	24.3	680	12	BI406237	157H03.Ma
38	388.6	24.3	720	12	BJ573369	BJ573369
39	388	24.2	672	14	CD323049	SS1_31_D1
40	383.2	23.9	621	13	BU635947	043C05.In
41	382.4	23.9	685	13	CA133117	SCSQRT103
42	380.4	23.7	718	14	CA801290	SAU03F04
43	380	23.7	615	13	BQ850509	QHB12R03
44	380	23.7	655	13	CA077350	SCQSAW103
45	379.4	23.7	788	14	CA295025	SCQGLV101

## ALIGNMENTS

RESULT 1	AY103779	Zea mays	FCO126548	mrna	linear	HTC	16-OCT-2002
LOCUS	AY103779	Zea mays	FCO126548	mrna	linear	HTC	16-OCT-2002
DEFINITION	AY103779	Zea mays	FCO126548	mrna	linear	HTC	16-OCT-2002
ACCESSION	AY103779	Zea mays	FCO126548	mrna	linear	HTC	16-OCT-2002
VERSION	AY103779.1	GI:21206857					
KEYWORDS	HTC						
SOURCE	Zea mays						
ORGANISM	Zea mays						

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 2333)  
Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitesitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.  
Maize Mapping Project/Dupont Consensus Sequences for Design of Overgo Probes  
Unpublished (2002)  
Coe, E.H.  
Direct Submission  
Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA  
If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the

maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

## FEATURES

Source

1. 2333  
/organism="Zea mays"  
/mol\_type="mRNA"  
/db\_xref="MaizeDB:637372"  
/db\_xref="taxon:4577"  
/clone\_lib="Maize Mapping Project/DuPont Consensus Library"

/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed Dupont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

## ORIGIN

Query Match 45.7%; Score 732.8; DB 11; Length 2333;  
Best Local Similarity 69.8%; Pred. No. 1.5e-108;  
Matches 989; Conservative 0; Mismatches 427; Indels 0; Gaps 0;

QY 1 CACACCGTGTGATTAACCATGAGCTATGATGCTCGATCTCAAGCTTCTGATCTTGAC 60  
DB 488 CACTCCAAAGTGCCTCACTGTGAGTCCATGATGCTCGACTCCGCGCGCGACCTCGAC 547

QY 61 AAGAGAGAACTCTGAGATCTTCAATGCTTCCGCTCTTGAAGGAATGCTCTTG 120  
DB 548 AAGGAGAGACCGCGAGATCTGCTCTTCTGCTCTTCTTCAAGGGGAATCAGTCTTA 607

QY 121 GAACCTGGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
DB 608 GAACCTGGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 667

QY 181 ATTCCTGCTGATTTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
DB 668 CTGGCACTAGCTTTTGAAGTGTGATTAAGAGAACCAAGCATTAATGGGATCAC 727

QY 241 AAAATGTCAATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
DB 728 AAGACATTAACCTTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 787

QY 301 TCATTTGATGATATTTCTCAATTTGCTTACTCATGATCTCTTCTGCTGCTGCTGCTGCT 360  
DB 788 TCTGTGATCTGATATTTTCAATCTGCTTATTAATGATCTTCTGCTGCTGCTGCTGCT 847

QY 361 AATTTGCTGAGAAATCTTGAATGCTTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
DB 848 AAGCTTGTGGGAATATGTTAAATGTTAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 907

QY 421 TCTTTTTCATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
DB 908 TCATGTTTTCATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 967

QY 481 CTAGGTTTCTACATGAGGCTTCAAGAGTGTCTTCAAGAGTGTCTTCAAGAGTGTCTTCA 540  
DB 968 CCAAGTTTATACAGATTTTAAAGAGGCTTCAATTTGATCAAGATGAGGTTGCTG 1027

QY 541 TATAGCTCTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
DB 1028 TTGAACTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1087

QY 601 CAGAACCAAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
DB 1088 CAAACCAAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1147

QY 661 CAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
DB 1148 AGATTTCTGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1207

QY 721 GGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780

DB 1208 GGTGAAGCTTTTGTGAGCACTGTGGAATCGAGACTACAAAGGAATTTGTGGGCTGCTC 1267  
QY 781 GACTTTGAAGCTTGGCCAGAAAGTCTCGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 840  
DB 1268 GATCTTAAACCGGCGCAGAAAGTACTTGTGTCGGATGTGGAATTTGGAGCGCGGACTTT 1327

QY 841 TACATGGCGGAGACCTTTGATGTTGAGGTTGTTGAGTGTGATCTCTCGTTAATGATTT 900  
DB 1328 TACATGGCTGCAACTATGATGTCATGTTCTTGGTATGATCTTTCGGTGAACATGGTT 1387

QY 901 TCTTTGCTGCTTGAAGCTTCTTATTTGGCTTAAATGCTGTTGAGTGTGAGGTAGCAGAT 960  
DB 1388 TCATTTGCAATTTGAACGTCGCATTTGACGCAAGTCTCTGTTGAAATTCGAAGTCTGCT 1447

QY 961 TGCAACCAAGATAAACTACTACCTGATCTCTTTTGAATGATCTATATAGCGTGCACACCAT 1020  
DB 1448 TGCACCAACAAAGGATTACCCAGAAATAGTTTGTGAGCTCATCTACAGCGCTGACACCATC 1507

QY 1021 CTGCATTTACAGCAACAGCTGCTGTTTGTAGATCTTCTACAAATGTTGAGCCAGGA 1080  
DB 1508 CTTCATACATACAGCAACAGCTGCTGTTTGTGAGGCTTCTTCAATTTGAGCTAAAGCCCGC 1567

QY 1081 GGTAAAGTCTTAATCAGTGAATTAATCTGCAAGAAAGCTGCTCCACCTCAGCTTCAATTCGCC 1140  
DB 1568 GGTAAAGTCTTAATCAGTGAATTAATCTGCAAGAAAGCTGCTCCACCTCAGCTTCAATTCGCC 1627

QY 1141 GCTTACATTTAAGCAGAGGGGATGATGATCTCATGATGTAAGGATATATGGGAGATGCTT 1200  
DB 1628 GGTACATTTAAGCAGAGGGGATGATGATCTCATGATGTAAGGATATATGGGAGATGCTT 1687

QY 1201 AAGATGCTGATTTGTTGATGTTCTTGGCGAGGATAGAACTGAGCAGTTCATTCGAGTT 1260  
DB 1688 AAGATGCTGATTTGTTGATGTTCTTGGCGAGGATAGAACTGAGCAGTTCATTCGAGTT 1747

QY 1261 CTACGGAAGGAATAGAGACTGTTGGAAGCAAAAGATGTTGTTCAATGATGATTTCTCT 1320  
DB 1748 CTACGGAAGGAATAGAGACTGTTGGAAGCAAAAGATGTTGTTCAATGATGATTTCTCT 1807

QY 1321 GAGGAGGATTAATGATGATGATGTTGAGGTTGGAATGATGATGTTGCGAGGAGTCCCAAG 1380  
DB 1808 CAGGAGGATTAATGATGATGATGTTGAGGTTGGAATGATGATGTTGCGAGGAGTCCCAAG 1867

QY 1381 GGTGAGCAACGATGGGCTCTGTTGTTGCCAAGAAG 1416  
DB 1868 GGTGAGCAACGATGGGCTCTGTTGTTGCCAAGAAG 1903

## RESULT 2

BE231445  
LOCUS  
DEFINITION  
S98015 Suaeda salsa ZAP cDNA library Suaeda maritima subsp. salsa  
CDNA similar to c.elegans CDNA YK92B11.3, mRNA sequence.

## ACCESSION

BE231445

## VERSION

BE231445.1

## KEYWORDS

EST.

## SOURCE

Suaeda maritima subsp. salsa

## ORGANISM

Suaeda maritima subsp. salsa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales; Amaranthaceae; Suaeda.

## REFERENCE

1 (bases 1 to 713)

## AUTHORS

Zhang, L., Ma, C.L., Wang, P.P., Sun, Y.F., Zhao, Y.X. and Zhang, H.

## TITLE

Expressed sequence tags from a halophyte Suaeda salsa cDNA library

## JOURNAL

Unpublished (2000)

## COMMENT

Contact: Rui Zhang

Key Laboratory of Plant Stress Research

The Biology Department of Shandong Normal University

No. 88, Weihai East Road, Jinan, Shandong Province, 250014, PRC

Tel: (86)531-2980864

Fax: (86)531-2988954

Email: zhangh@sdu.edu.cn

## FEATURES

source

1. 713



```

/organism="Suaeda maritima subsp. salsa"
/mol_type="mRNA"
/sub_species="salsa"
/db_xref="taxon:126914"
/dev_stage="seedling"
/clone_lib="Suaeda salsa ZAP cDNA library"
/notes="Organ: serial part tissue; Vector: lambda zap;
Site 1: EcoRI; Site 2: XhoI; total RNA extraction from
NaCl(400mM) treated Suaeda salsa by RNAGENT kit(Promega);
mRNA isolation by MESSAGEMAKER kit(GIBCO BRL);
directional cDNA synthesis(EcoRI XhoI) by cDNA synthesis
kit(STRATAGEN); the ZAP express library by GigapackIII
Gold Cloning kit(STRATAGENE)."

```

```

ORIGIN
Query Match 40.9%; Score 654; DB 10; Length 713;
Best Local Similarity 98.4%; Pred. No. 1.2e-95;
Matches 701; Conservative 0; Mismatches 5; Indels 4; Gaps 4;
QY 241 AAAAAATGCAAGTTATGTCGTGATGTGACTTCTCCCACTCTCAGTTTCCACCAT 300
DB 1 AAAAAATGCAAGTTATGTCGTGATGTGACTTCTCCCACTCTCAGTTTCCACCAT 60
QY 301 TCATTGATGTGATTTCCCAATGTTTACTCATGTATCTTCTGATGAGAGGTGGA 360
DB 61 TCATTGATGTGATTTCCCAATGTTTACTCATGTATCTTCTGATGAGAGGTGGA 120
QY 361 AATTGGTTGAAGATGTTGAATGGTTGAAGCCAGGGGGTTACATTTCTTCAGAGAA 420
DB 121 GAATGGTTGAAGATGTTGAATGGTTGAAGCCAGGGGGTTATATTTCTTCAGAGAA 180
QY 421 TCTTGTTCCTCAATCTGGGATCACAAACGCAAAAGCAATCCCACTACCTCGTGA 480
DB 181 TCTTGTTCCTCAATCTGGGATCACAAACGCAAAAGCAATCCCACTACCTCGTGA 240
QY 481 CTTAGTTCTACATAGGCTTCAAGAGTGTCTATTGCAAGATGATCTGGAATCTCT 540
DB 241 CTTAGTTCTACATAGGCTTCAAGAGTGTCTATTGCAAGATGATCTGGAATCTCT 300
QY 541 TATGAGCTCTCCCTACTAGTCTCAATGTATTTGAGCTTATGTGAGAAACAGAAAAAC 600
DB 301 TATGAGCTCTCCCTACTAGTCTCAATGTATTTGAGCTTATGTGAGAAACAGAAAAAC 360
QY 601 CAGAACAGATATGTTGGTGGCAAAAGTTGATCTTCAAGATGATAGGGGTTCCAG 660
DB 361 CAGAACAGATATGTTGGTGGCAAAAGTTGATCTTCAAGATGATAGGGGTTCCAG 420
QY 661 CGATTTCTGATAGTACAGTCAAGTGTATAGCATTTCTGCGATATG-AGCGTGATT 719
DB 421 CGATTTCTGATAGTACAGTCAAGTGTATAGCATTTCTGCGATATGAGAGCGTATT 480
QY 720 TGGCCCTGTTATGTTAGCACTGGAGGATATGAACCAACCAAGAGTTTGTGCAATGCT 779
DB 481 TGGCCCTGTTATGTTAGCACTGGAGGATATGAACCAACCAAGAGTTTGTGCAATGCT 540
QY 780 GGAATTTGAGCTCGCCAGAGTCTCTGATGTTGGTGGTGGTGGTGGTGGTGGTGGT 839
DB 541 GGAATTTGAGCTCGCCAGAGTCTCTGATGTTGGTGGTGGTGGTGGTGGTGGTGGT 599
QY 840 TTACATGCGGAGACCTTTGATGTTGAGGTTGTTGG-ATTGATCTCTCCGTTAATATGA 898
DB 600 TTACATGCGGAGACCTTTGATGTTGAGGTTGTTGGATTTGATCTCTCCGTTAATATGA 659
QY 899 TTTCCCTTGGCC-TTGAGCGTTCTATTTGGGCTTAATGCTGTTGAGTT 947
DB 660 TTTCCCTTGGCC-TTGAGCGGTCTATTTGGGCTTAATGCTGTTGAGTT 709

```

```

RESULT 3
---AY105261. 1012 bp mRNA linear HTC 16-OCT-2002
LOCUS
DEFINITION
Zea mays P00129777 mRNA sequence.
ACCESSION
AY105261

```

# VERSION KEYWORDS SOURCE ORGANISM

AY105261.1 GI:21208339  
HTC.  
Zea mays  
Zea mays

# REFERENCE AUTHORS

1 (Bases 1 to 1012)  
Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,  
Athur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.

# TITLE JOURNAL AUTHORS

Maize Mapping Project/DuPont Consensus Sequences for Design of  
Overgo Probes  
Unpublished (2002)  
2 (Bases 1 to 1012)  
Coe, E.H.

# JOURNAL TITLE COMMENT

Submitted (25-APR-2002) Maize Mapping Project, University of  
Missouri, Columbia, MO 65211, USA  
If you are interested in getting corresponding physical clones,  
these are publicly available from ZmDB and may be found by BLAST  
searching at MSL, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR,  
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the  
maize cDNA sequences is either Virginia Walbot, Stanford or Pat  
Schnable, Iowa State, then clones may be requested from ZmDB:  
www.zmdb.iastate.edu.

# FEATURES source

1..1012  
/organism="Zea mays"  
/mol\_type="mRNA"  
/db\_xref="MaizeDB:637558"  
/db\_xref="taxon:4577"  
/clone\_lib="Maize Mapping Project/DuPont Consensus  
Library"  
/note="this sequence is part of a project of EST  
assemblies resulting from the application of public  
contigs to seed Dupont contigs; this resource was  
assembled by DuPont as part of a collaboration for the  
overgo addressing of BACs in conjunction with the Maize  
Mapping Project"

# ORIGIN

```

Query Match 35.4%; Score 566.6; DB 11; Length 1012;
Best Local Similarity 73.2%; Pred. No. 1.2e-81;
Matches 739; Conservative 0; Mismatches 270; Indels 1; Gaps 1;
QY 311 TGATATTTCCAAATGTTTACTCATGTATCTTTCTGATGAGAGGTGAAATTTGGTTG 370
DB 2 TGATATTTCCAAATGTTTACTCATGTATCTTTCTGATGAGAGGTGAAATTTGGTTG 61
QY 371 AAAGATGTTTCAAAATGTTTCAAGCCAGGGGTTTACATTTCTTCAGAGATCTTTGTTTC 430
DB 62 AGAGATGTTTCAAAATGTTTCAAGGTTTGGTGGCTATATCTTTTAGGCAATCTTGTTC 121
QY 431 ATCAATCTGGGATCACAAAGCAATCCCACTACCTACCTACCTACCTACCTACCTACCT 490
DB 122 ATCAATCTGGAGATTCAAAAGGAAAGTTAATCCGACACATATAGGGAACCAAGTTT 181
QY 491 ACATTAAGGCTTCAAAAGGTGTCAATTTGCAAGATGATCTGGAACCTCTTATGAGCTCT 550
DB 182 ATACTAAGGTTTCAAAAGATGCCATACCTTTTGTATCAAGATGGGAATCTTTCGAATCT 241
QY 551 CCCTACTTACTCTCAATGTATTGGAGCTTATGTCAGAAACCAAGAAACCAAGAACCA 610
DB 242 CTCCTGTTACTTGCAGTGTATTGGTCTTATGTTAAACCAAGAAACCAAGAACCA 301
QY 611 TTAGTGTGTTGGCAAAAGTTGATTTCTAAGATGATGAGGGGTTTCAGCGATTTCTGG 670
DB 302 TATGTTGGCTATGCAAAAGGTCATTTCTACAGAGATAAGGATTTCAAGATTTTGG 361
QY 671 "ATACTAGCAGTACAAATGTAAATAGCTT-CTGCGATATGAGCGTATTTTGGCCCTG 729
DB 362 ACAATGTCAGTACAAAGCCAGTGGGANTATTACGTTACGAGCGGCTTTTGGAGAGGT 421

```



```

QY 730 TATGTTAGCACTGGAGGATATGAACCAACCAAGAGTTTGTGTCAATGCTGGACTGAAG 789
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
422 TATGTGAGCACTGGTGGAGTTGAGACTACAAAGAAATTTGTGCAAGCTGGATCTCAA 481
QY 790 CTGGCCAGAGGTCCTGGAGTTGTTGTTGGAAATTTGGGAGGTGACTTTTACATGGC 849
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
482 CTGGACATAGGTGCTTGATTTGATTTGATTTGATTTGGGAGGTGACTTTTATATGGC 541
QY 850 GAGACCTTTGATTTGAGTTGTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 909
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
542 GAAATAATGATGCTCATGTTGTTGGTATTTGATTTTCCATAAAGATGATCATTTGCA 601
QY 910 CTTGAGCCTTCTATTTGGGCTTAAATGTGCTGTTGAGTTTTCAGGTAGCAGATTCACCAAG 969
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
602 CTTGAGCCTGCAATGGGCGAGTTGCTCAGTGGAGTTTGAAGTTGCTGATTGCACTACG 661
QY 970 ATAACTACCTGATTAATCTTTTGTGATCATCTATAGCGGTGAGCACCATTCTGATATT 1029
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
662 AGACATACCCAGACACACATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 721
QY 1030 CAGGACAGCTGGTGTGTTTGTAGATCTTCTACAAATGTTGAGCGAGGATTAAGTT 1089
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
722 CATGACAAACCTCTTTTGTCAAAAGTTTTCAGTGGCTGAGGCTGGGGGCAAGTC 781
QY 1090 CTAATCAGTGTATCTGACAGAAAGCTGGTCCACCTCAGCTGAATTCGCCGCTTACAT 1149
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
782 CTTATCAGTACTACTGCGAGTCTCTGGGAACCATCAGAGGATTTGCGCGTACAT 841
QY 1150 AAGCAGAGGATGATGATCTCCATGATGTAAGGAATATGGGAGATGCTTAAGATGCT 1209
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
842 AAGCAGAGGATGATGATCTCCATGATGTAAGGAATATGGGAGATGCTTAAGATGCT 901
QY 1210 GGAATTTGTTGATTTCTTGGGAGGATGAGTCTGAGGATTTCAATCGATTTCTACGAG 1269
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
902 GGTTTTCTGATGATCTGCTGAGGATCGAAGTCAATGATGCTGCTGCTGCTGCTGCT 961
QY 1270 GAACATAGAGCTGTTGAGAGGAAAGAGTGTGTTCAATGATGATTTCTC 1319
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
962 GAGTTAGCTGAATTTGAAGAAACAGAGATTTCTGCTGACTTCTAC 1011

```

```

RESULT 4
BS574969
LOCUS
DEFINITION
E07 Triphysaria versicolor root-tip, early DMO-induced transcript
BS574969
VERSION
KEYWORDS
SOURCE
ORGANISM
Triphysaria versicolor
Triphysaria versicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; Lamiales; Orobanchaceae; Rhinanthaceae;
Triphysaria.
1 (bases 1 to 1211)
Matvienko, M., Torres, M.J. and Yoder, J.I.
Transcriptional responses in the hemiparasitic plant Triphysaria
versicolor to host plant signals
Plant Physiol. 127 (1), 272-282 (2001)
21437952
11553755
Contact: Yoder, J.I.
John I. Yoder Research Lab, Dept. of Vegetable Crops
University of California at Davis
137 Armadillo Hall, One Shields Drive, Davis, CA 95616, USA
Tel: 530 752 1741
Fax: 530 752 9659
Email: jyoder@ucdavis.edu
length= 1211 bp, Location/Qualifiers
i. .1211
/organism="Triphysaria versicolor"

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

```

```

FEATURES
source

```

```

/mol_type="mrna"
/db_xref="taxon:64093"
/tissue_type="root-tips ~5mm length"
/dev_stage="3-4 weeks growth"
/lab_host="E. coli"
/clone_lib="Triphysaria versicolor root-tip, early
DMO-induced transcript cDNA library"
/notes="vector: PCR2.1 TA Cloning System, Invitrogen,
Carlsbad, CA; PCR-based suppression subtractive
hybridization cDNA library"

```

## ORIGIN

```

Query Match      33.6%; Score 538.6; DB 10; Length 1211;
Best Local Similarity .72.8%; Pred. No. 3.4e-77;
Matches 694; Conservative 0; Mismatches 259; Indels 0; Gaps 0;

QY 1 CACACGGTGTGATTAACCATTTGAGCTATGATGCTGCTGATTTCTCAAGCTTCTGATCTTGCAC 60
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
258 CATTCTGTGTACCTGACCTGGATTCATGATGCTGATTTCCAAAGCATCTGATCTCGAT 317
QY 61 AAAGAAGAACTCTCTGAGATTTCTTCAATGCTTCGSCCTCTTGAAGGAAATGCTCTTTG 120
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
318 AAGGAAGAAAGGCTCTGAGTGTATCTATGCTCCCAACCATATGAAGGCAAGTCGCTTCTT 377
QY 121 GAACTTGGGGCTGCTGATTTGCTGTTTACTGTGGAATGGCTGAGAAAGCTGGCCAGGTT 180
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
378 GAACTGGGAGCTGCTGATTTGCGCTTCTACTGCTGATTTAGCTAAGAGGCCAAGAGTT 437
QY 181 ATTGCTCTGATTTTCATGAGAGTGTCTATCAGAGCAATTAAGTATCAATGGGCACTAC 240
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
438 GTAGCTTTAGACTTCTATCGAAGCATATTAAGAGAGATGAAGCCCTCAATGGCATCAT 497
QY 241 AAAAATGTCAAGTTTATGTTGCTGATGAGTCTTCTCCACTCTCAGTTTCCACACCAT 300
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
498 AAAAATGTCAAAATTCATTTGCTGATGTCACATCCCGAGACTTGAACCTTCTCTGAAGA 557
QY 301 TCATTTGAGATGATATTTCTCCAAATTTGTTACTCATGATGTTCTTCTGATGAAGGTGGA 360
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
558 TCATTTGAGTGTATTTCTGACTGCTGATTTGATTTGATTTGATTTGATTTGATTTGAT 617
QY 361 AATTTGTTGAAAGATTTGAAATGTTGAAGCTTGAAGCCAGGGGTTTACATTTTCTTCAAGAA 420
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
618 AATTTGAGAGAGAGAGTGTCTCAATGTTGAAAGTTGTTGGCTTACATATATTTTCAAGAA 677
QY 421 TCTTGTTCCTCAATCTGGGATCACAACGCAAAAGCAATCCACCACTACCTGCTGAA 480
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
678 TCATGTTTCCATCAGTCTGCTGACCAACAAAGAAAGATTAATTCGACTCATTTATCGCGAA 737
QY 481 CCTAGGTTCTACACTAAGGCTTCAAGAGTGTCTTTCAGAGATGATCTGGAAGTCTT 540
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
738 CCAAGATTTTATACCAAGTTGTTCAAGAAATGCCATATGATGATTAAGTCTTGAATTTCA 797
QY 541 TATGAGCTCTCCTACTTACTGCAATGTTATTTGAGCTTATGTCAGAAACCAAGAAAC 600
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
798 TTTGAACCTTGCTTATTTGTTGTTGATGCTGCTGCTTATGTTGATGCTTATGTTGATGCT 857
QY 601 CAGAACCAAGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 660
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
858 CAGATCAGATCTCTGCTGATTTGCAAGAGTGTAGCTCAGATCATGACAGAGATTTCCAG 917
QY 661 CGATTTCTGATATAGCCAGTACAGTGTAAATAGATTTCTGCGATATGAGGCTGTATT 720
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
918 AAAATTTCTGGATTAATGTTCCAAATATTAATGTTAAATGTTAAATGTTAAATGTTAAAT 977
QY 721 GGCCTCTGTTATGTTTGTAGCACTGGAGGATATGAACCAACCAAGAGTGTGCTCAATGCTG 780
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
978 GCAAGGAGATATGAGGACCGGTGGAGTGAAGAAACAACTAAAGATTTTGTAGCCAACTG 1037
QY 781 GACTTGAAGCTGCGCAGAGGCTCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 840
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1038 GACTTGAAGCTGCGCAGAGGCTCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1097
QY 841 TACATGCGGAGAGCTTTGATGTTGAGGTTGTTGATTTGATCTCTCCGTTAATATGATT 900

```

Db 1098 TACATGGCTGACGACGATGATGTTTCATGTTGTTGGCATTGACCTCTCGCTCAATATGATT 1157  
 Qy 901 TCTTTGGCTTGAGCGTTCATGAGCTTAAATGTCCTGCTTCAAGTTGAGCT 953  
 Db 1158 TCTTTGCTCTCGAGCGTGCATGTTGCTCTCAGTGTGCTGTGAGTTGAGTT 1210

RESULT 5  
 CK206655 1022 bp mRNA linear EST 08-DEC-2003  
 LOCUS FGAS018258 Triticum aestivum FGAS: Library 5 GATE 7 Triticum  
 DEFINITION aestivum cDNA, mRNA sequence.  
 ACCESSION CK206655  
 VERSION CK206655.1 GI:39569045  
 KEYWORDS EST.  
 SOURCE Triticum aestivum (bread wheat)  
 ORGANISM Triticum aestivum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Poideae; Triticeae; Triticum.  
 1 (bases 1 to 1022)  
 AUTHORS Allard, F., Crosby, W.L., Danyluk, J., Eudes, F., Frick, M., Gaudet, D.,  
 Genswein, B., Graf, R., Gulick, P., Hyman, L.D., Larocque, A.,  
 Links, M.G., McCarthy, B.L., Monroy, A., Muzak, I., Nilsson, D.,  
 Penniket, C., Roach, J.L. and Sarhan, P.  
 TITLE Functional Genomics of Abiotic Stress in Wheat and Canola Crops  
 JOURNAL Unpublished (2003)  
 COMMENT Contact: Mm L Crosby  
 Bioinformatics  
 University of Saskatchewan, Department of Computer Science  
 1C101 Engineering Building, 57 Campus Drive, Saskatoon,  
 Saskatchewan, S7N 5A9, Canada  
 Tel: 306 966 1769  
 Fax: 306 966 2033  
 Email: fgas.est@usask.ca

This sequence is the direct result of the Base calling software  
 Phred (default parameters). It is the raw base calls. To aid in the  
 identification of the high quality insert the software Lucy  
 (default parameters) has been run on this sequence. Lucy identified  
 the region [9,84].  
 Plate: L58008 row: C column: 07.  
 Location/Qualifiers  
 1..1022  
 /organism="Triticum aestivum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4565"  
 /clone\_lib="Triticum aestivum FGAS: Library 5 GATE 7"  
 /notes="Vector: pcwv.sp08; Crown and developmental stages  
 of spike formation in wheat cultivar Norstar. 4 mRNA  
 populations were combined before constructing the library.  
 The first mRNA population is from 1cm crown sections after  
 30 days of cold acclimation. The second is from 1cm crown  
 sections after 11 days of deacclimation (before  
 deacclimation plants were fully vernalized for 49 days).  
 The third is from different developmental stages of spike  
 formation (5 to 50mm) that still have not emerged from the  
 leaf (dissection required). The last is from different  
 developmental stages of spike and seed formation after  
 having emerged from the leaf (visible). First strand  
 synthesis in this library was done in the presence of  
 methylated dCTP thereby protecting from internal cleavage  
 with NotI."

## FEATURES

source

## ORIGIN

Query Match 29.1%; Score 465.6; DB 14; Length 1022;  
 Best Local Similarity 70.0%; Pred. No. 2e-65;  
 Matches 627; Conservative 0; Mismatches 269; Indels 0; Gaps 0;

Qy 521 AAGATGGATCTCGAACTCTTATGAGCTCTCCCTACTAGTCGCAATGATGAGCTT 580  
 Db 25 ATGACCAAGATGGCAATTCCTTTGAGCTTCTCTGCAACTTCCCAAGTCATTGAGCTT 84

Qy 581 ATGTCAGAAACAGAAACCCAGAACCGAGTATTAGTTGGTCTGGCCAAAAGTTGATTCTA 640  
 Db 85 ATGTGAGAGCAGAGAACCCAGAACCCAGATATGTTGGCTATATGGGAGAGGTCAGATGCA 144  
 Qy 641 AGATGATPAGGGGTTCCAGCGATTTCTGGATACTAGCCAGTACAAGTGTAAATAGCATTC 700  
 Db 145 CAGAAGACAAAGGCTTTTCAGAGATTCTGACAAATGTGCAATACAAATCCATGGAATCC 204  
 Qy 701 TGGCATATGAGCTGTATTTTGGCCCTGTTATGTATAGCACTGGAGGATATGAACCAACA 760  
 Db 205 TGGGTTATGACGGCTGTTTGGGNGGGTTATGTACAGACCGTGGATTCGAGACACAA 264  
 Qy 761 AAGAGTTTGTGTCATGCTGAGCTTGAAGCTGCGCAGAGGCTCTGAGATTTGTTGTG 820  
 Db 265 AAGAATTTGTGACAGCTGTGATCTGAAGCTGCGCAGAGGCTCTGAGATTTGTTGTG 324  
 Qy 821 GAATTTGGAGGTGATCTTTTACATGCGGAGACCTTTTGTATGTTGAGGTTGTTGATTTG 880  
 Db 325 GCATCGGAGGAGGTGACTTTCTATGCTGAAACCTATGATGTCATGCTCTCGGCATCG 384  
 Qy 881 ATCTCTCCGTTAAATGATTTTCTTTGCGCTTGAGGCTTCTATGCGGCTTAAATGCTG 940  
 Db 395 ATCTTCCATCAACATGTTTGTTCGATCGAGGCTGCCATCGGCGCTCGTCTCGG 444  
 Qy 941 TTGATTTGAGGTAGCAGATGCAACCAAGATAAATACCTGATTAATCTTTTGTGATGTC 1000  
 Db 445 TTGAGTTTGAAGTTGCTGACTGCAACCAACCAAGATAACGAGAACACAGTTTGTATGA 504  
 Qy 1001 TCTATAGCCGTGACACCAATCTGCAATTCAGGACAAAGCTCGCTTGTGTAGTCTTCT 1060  
 Db 505 TCTACAGCCGTGACACCAATCTCTCCATTCAGACAAACCTCTCTGTTGAGAACTTCT 564  
 Qy 1061 ACMAATGTTGAGCCAGGAGGTAAAGTTCTAATCAGTATTTACTGCAAGAAAGCTGCTC 1120  
 Db 565 TCAAGTGGCTCAAGCTGGGGGCAAGTGTCTATCAGTGACTACTGCGAGGAGCCCTGGA 624  
 Qy 1121 CACCTCACTGAATTCGCGCTTACATTAAGCAGAGGGGATATGATCTCCATGATGTA 1180  
 Db 625 CACCTCAGAGGAATTTGCTGATATACATCAGCAGAGAGGCTATGACCTCCATGACGTA 684  
 Qy 1181 AGGAATATGGCAGATGCTTAAAGATGCTGGATTTGTTGATGTTCTTCCCGAGTAGAA 1240  
 Db 685 AGACCTATGGAAGATGCTGAGATGCGGTTTCCAGTGTCTGCGGAGAGACCGCA 744  
 Qy 1241 CTGAGCAGTTCATTCGAGTTCACGGAAGAACTAGACACTGTTGAGAGGAAAAGGATG 1300  
 Db 745 CCGACCAAGTTCTGACGCTCTCGAGAGGAGCTGGCGAGAACGAGAAACAGGAGG 804  
 Qy 1301 TGTTCATTAGTATTTCTCTGAGGAGGATTTACAATGACATTTGAGGTTGGATGATA 1360  
 Db 805 CTTCTCGGGGACTTCCACCCAGAGGACTACGACGACATGTCACCGCTGGAGCGGA 864  
 Qy 1361 AGTTGCGGAGCTGCGCAAGGTGAGCAACGATGGGCTCTGTTCTGTTCCCAAGAG 1416  
 Db 865 AGCTGAAGCGAGCTCCCGCGGAGCAGAGTGGGGGCTGTTTCATCGGAACAGG 920

## RESULT 6

BM408047

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

BM408047 767 bp mRNA linear EST 10-MAR-2003  
 EST582374 potato roots Solanum tuberosum cDNA clone cPR033G1 5'  
 end, mRNA sequence.  
 BM408047  
 BM408047.1 GI:18259677  
 EST.  
 Solanum tuberosum (potato)  
 Solanum tuberosum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;  
 Asteridae; Limnids; Solanales; Solanaceae; Solanum.  
 1 (bases 1 to 767)  
 van der Hoeven, R., Sun, H., Karameycheva, S.A., Tsai, J., Van Aken, S.,  
 Utterback, T., Chiemingo, A., Bougri, O., Buell, C.R., Renning, C.,

Tankley, S. and Baker, B.  
 Generation of ESTs from potato roots  
 Unpublished (2001)  
 Contact: Robin Buell  
 The Institute for Genomic Research  
 9712 Medical Center Dr, Rockville, MD 20850, USA  
 Email: potato-array@tigr.org  
 This clone can be obtained from the University of Arizona Genomics  
 Institute. Orders can be made through URL:  
 http://genome.arizona.edu/orders/  
 Seq primer: T3.

#### FEATURES

Location/Qualifiers  
 1..767  
 /organism="Solanum tuberosum"  
 /mol\_type="mRNA"  
 /cultivar="Kennebec"  
 /db\_xref="taxon:4113"  
 /clone="CPRO33G1"  
 /tissue\_type="roots"  
 /dev\_stage="in vitro grown stem cuttings"  
 /lab\_host="SOLR"  
 /clone\_lib="potato roots"  
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
 xhoI; supplier: Cornell University, Tankley lab;  
 sequencing: The Institute for Genomic Research. Roots were  
 isolated from in vitro grown stem cuttings on CM medium.  
 roots were isolated two weeks after placing the stem  
 cuttings from in vitro grown plants on medium."

#### ORIGIN

Query Match 29.0%; Score 464.6; DB 12; Length 767;  
 Best Local Similarity 75.4%; Pred. No. 3.3e-65;  
 Matches 578; Conservative 0; Mismatches 189; Indels 0; Gaps 0;

QY 439 GGGGATCACAAGCGAAGCAATCCACACCTACCGTCACTAGGTTCTACACTAG 498  
 DB 1 GGGACCAACAGGAAAGAACCCCAACCATTTATCGGAGCTAGATTACACAA 60  
 QY 499 GCCTTCAAAGATGTCATTGCAAGATGATCTCGAACTCTTATGAGCTCTCCCTACT 558  
 DB 61 GTCTTTAAAGATGTCATATAAATGCTGTGATGTAATCATTTGAACITTTCTCTCAT 120  
 QY 559 AGCTGCAATGTTAGGAGCTTATGTGCAAGAACCAAGAAACCAAGACAGATTAGTGG 618  
 DB 121 GGTTCAGTGCATTGGCTTATGTTAAACCAAAAGAAATCAGAAATGATTTGTTGG 180  
 QY 619 TTGTGCAAAAGTTGATCTTAAGCATATAGGCTTCCAGGATTCGAGATTCGATAC 678  
 DB 181 CTATGCAAAAGGTTAAATCTGAGGATCAGAGGGAATCCAGCGTTTCTCGGCAATGTT 240  
 QY 679 CAGTACAGTGAATAGCAATCTCGCATATGAGCGTGTATTTGGCCCTGTTATGTAGC 738  
 DB 241 CAATACAAATGTAGTGGCATCTGCGATATGAACGTGTCTTTGGAGAAAGTTATGTGAGC 300  
 QY 739 ACTGGAGGATATGAACCAACCAAGATTTGTGTCAATGCTGGAATTTGAAGCTCGCCAG 798  
 DB 301 ACAGAGGACTGTATACCAACAAGAAATGCTTTATGTGATCTTCAACCTGGCCAA 360  
 QY 799 AAGTCTGAGATGTTGGTTGTGGAATTTGTGTGAGGTGATCTTTACATGGCGGACCTTT 858  
 DB 361 AATGCTCTGATGTTGGCTGTGGAATAGGTGAGGTGATCTTTACATGGCTGAGAAATAT 420  
 QY 859 GATGTTGAGATTTCTGGATTTGATCTCTCCGTTATATGATTTCTTTTGGCCCTTGAGCGT 918  
 DB 421 GATGTTCTATGTTGTTGCTATTCACCTCTCGATTACATGATATGTTTGTCTTGAACGT 480  
 QY 919 TCTATTGGGCTTAATGCTGCTGCTGATTTGAGTGGATGAGATTCACCAAGATAAATAC 978  
 DB 481 TCAATGCTCTCAATGTCGCTGATTTGAGTTGCTGATTTGATTTGATTTGATTTGATTT 540  
 QY 979 CCTGATTAATCTTTTGAATGTCATCTATAGCGGTGACCAATCTTGCATATTCAGGACAAG 1038  
 DB 541 CCTGATGGCAATTTGATGTGATTTTACAGTCGAGACATATCTCTTCATCTCAGACAA 600

QY 1039 CTGTGGTTGTTTACATCTCTTACAAATGTTTCAAGCCAGAGGTTAAAGTTCTTAATCAGT 1098  
 DB 601 CCCAGTTATTACAGATCTTTCTACAGTGGCTGNAACAGGAGGCAAGTCTCTAATAGT 660  
 QY 1099 GATTACTGCAAGAAAGCTGGTCCACCTACCTGATTCGCGCTTACATTAAGCAGAGG 1158  
 DB 661 GATTACTGCAATATGCCAATACCAAGCATCGAGGATTTCTGATGATATAAGCAAGG 720  
 QY 1159 GGATATGATCTCCATGATGTAAAGGAATATGGCAGATGCTTAAAGA 1205  
 DB 721 GGTATGATTTACATGATGTTTCAACATATGGCCAGATGCTCAAGA 767

#### RESULT 7

CF208434 781 bp mRNA linear EST 01-AUG-2003  
 CAB20003\_Ila\_Fa\_G05 Cabernet Sauvignon Flower bloom - CAB2 Vitis  
 vinifera cDNA clone CAB20003\_Ila\_Fa\_G05 5', mRNA sequence.  
 CF208434  
 CF208434.1 GI:33402807

#### ACCESSION

#### VERSION

#### KEYWORDS

#### SOURCE

#### ORGANISM

#### REFERENCE

#### AUTHORS

#### TITLE

#### JOURNAL

#### COMMENT

#### FEATURES

#### source

#### Location/Qualifiers

#### 1..781

#### /organism="Vitis vinifera"

#### /mol\_type="mRNA"

#### /cultivar="Cabernet Sauvignon"

#### /db\_xref="taxon:29760"

#### /clone="CAB20003\_Ila\_Fa\_G05"

#### /sex="Hermaphrodite"

#### /dev\_stage="Bloom"

#### /clone\_lib="Cabernet Sauvignon Flower bloom - CAB2"

#### /note="Organ: Flower - Bloom; Vector: pDNR; Site 1: SfiI; Site 2: SfiI; CAB2 is a cDNA library of Vitis vinifera cv. 'Cabernet Sauvignon' Clone 8 berries. Samples were collected at full bloom (80 to 100% flowers showing dehiscence of calyxtraps or caps and anthers fully extended). Sampled vines were located at the University of California, Davis, Experimental Vineyard. cDNAs were made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows: 5'-AAGCACTGTTATCAACGAGATGCTCCATACGCGCGG-3' and 5'-ATTAGAGCGGCGGCGGCGGATG-3' (30'NN-3'. Library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-3 kb size fraction."

#### ORIGIN

#### Query Match

#### Best Local Similarity

#### Matches

#### Conservative

#### Mismatches

#### Indels

#### Gaps

#### 0;

#### 0;

#### 0;

#### 0;

#### 0;

#### 0;

#### 0;

#### 0;



613 TCTGAATCTCAGTCATCCTCAATGTTGGAGCTTATGTGAGAAACAAAGAAT 672  
 601 CAGAACAGATTAGTGTGTCGCAAAAGTTGATCTTAAGCATGTAAGGGTTCAG 660  
 673 CAAACCAAGTATGTTATGTCAGATGCTTAATTCACAAATGACAAAGGGTTCGA 732  
 661 CGATTTCTGGATACACGAGTACAGTGAATAGCATCTCGGATATGAGGCTGATTT 720  
 733 CGTTCTCTGGATATGACAGTAAATGACGGGCTATATACGCTATGAAAGTGTCTTT 792  
 721 GGCCTCTGTGTATGACACCTGAGGATATGAACACCAACCAAGAGTTGTGTCATGCTG 780  
 793 GGTGAAGGCTTGTGTGATACAGGGGACTTGAACCAACAAAGAAATTTGTGACAAATG 852  
 781 GACTTGAAGCTGGCCA 797  
 853 GACCTTAAGCCTAGCCA 869

RESULT 9  
 CK166076/c  
 LOCUS  
 DEFINITION 1115 bp mRNA linear EST 05-DEC-2003  
 FGAS050129 Triticum aestivum FGAS: Talt7 Triticum aestivum cDNA,  
 mRNA sequence.

ACCESSION CK166076  
 VERSION CK166076.1 GI:38998769

KEYWORDS EST  
 SOURCE Triticum aestivum (bread wheat)  
 ORGANISM Triticum aestivum

REFERENCE  
 AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Poideae; Triticeae; Triticum.

1 (bases 1 to 1115)  
 Allard, F., Crosby, W.L., Danyluk, J., Budes, P., Frick, M., Gaudet, D.,  
 Genswein, B., Graf, R., Gulick, P., Hrycan, L.D., Laroché, A.,  
 Linka, M.G., McCarthy, E.L., Monroy, A., Muzak, I., Nilsson, D.,  
 Penikett, C., Roach, J.L. and Sarhan, F.,  
 Functional Genomics of Abiotic Stress in Wheat and Canola Crops  
 Unpublished (2003)  
 Contact: Wm L Crosby  
 Bioinformatics  
 University of Saskatchewan, Department of Computer Science  
 1C101 Engineering Building, 57 Campus Drive, Saskatoon,  
 Saskatchewan, S7N 5A9, Canada  
 Tel: 306 966 1769  
 Fax: 306 966 2033  
 Email: fgas\_estseqs.usask.ca

This sequence is the direct result of the Base calling software  
 Phred (default parameters). It is the raw base calls. To aid in the  
 identification of the high quality insert the software Lucy  
 (default parameters) has been run on this sequence. Lucy identified  
 the region [65,863].  
 Plate: Talt707 row: K column: 02.

#### FEATURES

Source  
 1..1115  
 /organism="Triticum aestivum"  
 /mol\_type="mRNA"  
 /cultivar="Wheat line CI 14106"  
 /db\_xref="taxon:4565"  
 /lab\_host="DH5 alpha"  
 /clone\_lib="Triticum aestivum FGAS: Talt7"  
 /note="Organ: Crown; Vector: pGEM-T; SSH (suppression  
 subtractive hybridization) cDNA library from genotype  
 CI14106 cold hardened at 2 C for 21 days and 49 days  
 (equal amount of cDNA pooled together before subtraction,  
 teater) and subtracted against genotype CI14106  
 non-hardened (20 C) (driver). Nitro-pyrole anchored  
 oligo-dt priming and non-directional cloning."

#### ORIGIN

Query Match 28.2%; Score 452; DB 14; Length 1115;  
 Best Local Similarity 68.9%; Pred. No. 2.9e-63;

	Matches	663;	Conservative	0;	Mismatches	295;	Indels	4;	Gaps	3;
Qy	455	AAAGCAATCCCAACCCACTACCGTGAACCTAGGTTCTACACTAAGGCTTCAAGAGTGTG	514							
Db	1103	AAAGTAACCCGACCCACTACCGGAGCGAGGTTTACACCAAGGTGTAAAGGA--TGC	1046							
Qy	515	ATTTCGAAGATGATCTGGAACCTCTTATGAGCTCTCCCTACTAGCTGCAATGATTTG	574							
Db	1045	CATCTCTAGACCAAGAGGGATTCCTTGAGCT--TCCTTCGTAACCTCCAGTGCATG	987							
Qy	575	GAGCTTATGTCAGAAAACAAGAAAACAGAACCCAGATAGTTGGTTGGGCAAAAGTTG	634							
Db	986	GAGCTTATGTCAGAAAACAGAGGAGAACCCAGACCCAGATATGTTGGCTGTGGGAAGG	927							
Qy	635	ATTCTAAGATGATTAAGCGGTTCCAGCGATTTCTCGATACTAGCCAGTACCAAGTGA	694							
Db	926	AGTCACAGAGAACAAAGGCTTTCAGAGATCCCTGCAATGTGCAGTACAATCCACTG	867							
Qy	695	GCATCTCGATATGAGCGGTATTTGGCCCTGTTATGTTAGCACTGGAGGATGAA	754							
Db	866	GAATCTTGGTATGAGCGGT--GTTGGGAGGGTATGTTAGCACCGGTGATTCGAGA	808							
Qy	755	CCACAAAGAGTTTGTGTCAATGCTGACTGAAGCTGGCCAGAGGTCCTGGAGTTG	814							
Db	807	CCACAAAGAGTTTGTGTGCAAAAGCTGCAAGCTGCTCAGAAAGTGTCTCGATGAG	748							
Qy	815	GTGTGGAATTTGGTGGAGGTGACTTTTACATGCGCGAGACCTTTGATGTTGAGGTT	874							
Db	747	GGTGTGATTCGAGAGGGGACTTCTACATGGCTGAACCTATGATGATTCCTGCTCG	688							
Qy	875	GAATTCATCTCCGTTAAATGATTTCTTTCGCTTGAGCGTCTTATTTGGGCTTAAAT	934							
Db	687	GCATCGATCTTCTTATCAACATGGTTTGGTTCCAAATGAGCGTGCATCGGGCGCTG	628							
Qy	935	GTGCTGTGAGTTTGAAGTACAGATGTCACCAAGATAAACTACCTCGATTAATCTTTG	994							
Db	627	GCTCGCTTGGTTGAGGTTGCTGACTGCAACCAAGGAATACGACAGAGAACACGCTT	568							
Qy	995	ATGTCATCTATAGCGGTGACACCACTCTGCATATTCAGGACAGCGCTGTTGTTAGAT	1054							
Db	567	ATGTCATCTATAGCGGTGACACCACTCTGCATATTCAGGACAGCGCTGTTGTTAGAT	508							
Qy	1055	CCTTCTACAAATGGTTGAAGCCAGAGGTAAAGTTCTAATCAGTGAATTAATTCGAAGA	1114							
Db	507	ACTTCTTCAAGTGGCTCAAGCTCGGGGCAAGTGTGATCACTGATCTACGAGGAGCC	448							
Qy	1115	CTGCTCCACCTCCCTCACTGAAATTCGCGCTTACATTAAAGCAGAGGGATATGATCTCA	1174							
Db	447	CTGGGACACCATCAGAGGAATTCGCTGTGTACATCAAGCAGAGAGGCTATGACCTCAT	388							
Qy	1175	ATGTAAGGATATGCGGAGATGCTTAAAGATGCTGAATTTGATGTTCTTCCGAGG	1234							
Db	387	ATGTAAGGATATGCGGAGATGCTTAAAGATGCTGAATTTGATGTTCTTCCGAGG	328							
Qy	1235	ATAGAATCTGAGCAGTTCATTCGAGTTCCTACGGAAGGAATCTAGAGACTGTTGAGA	1294							
Db	327	ACCGCACCGACCCAGTTCCTGAGGCTCTCGAGAGAGGAGCTGGGCGAGACCGAGA	268							
Qy	1295	AGATGCTGCTAATAGTATTCCTGAGGAGGATTAACATGATGTTGGAGTTGCA	1354							
Db	267	AGAGGCTCTTCTGCGGACTTCACCCAGAGGAGTACGAGGATGCTACACGCGCGA	208							
Qy	1355	ATGATAAGTTGCGGAGGCTCCCAAGGCTGAGCAACGATGGGTCTGTTGTCGCAAGA	1414							
Db	207	GCGCGAAGTTGAAGCGGAGCTCTGCCCGGAGCAGAAAGTGGGGGCTGTTTCATCG	148							
Qy	1415	AG 1416								
Db	147	AG 146								

RESULT 10  
 CB970167

LOCUS CB970167 753 bp mRNA linear EST 30-APR-2003  
 DEFINITION CAB10003\_11a\_Fa\_H06 Cabernet Sauvignon Flower Pre-bloom - CAB1  
 Vitis vinifera cDNA clone CAB10003\_11a\_Fa\_H06 5', mRNA sequence.  
 ACCESSION CB970167  
 VERSION CB970167.1 GI:30253616  
 KEYWORDS EST.  
 SOURCE Vitis vinifera  
 ORGANISM Vitis vinifera  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; Vitaceae; Vitis.  
 1 (bases 1 to 753)  
 Goes da Silva, P., Iandolino, A., Lim, H., Baek, J., Jones, K. and  
 Cook, D.  
 Expressed sequence tags from Vitis vinifera 'Cabernet Sauvignon'  
 berries at various developmental stages  
 Unpublished (2003)  
 JOURNAL Contact: Douglas Cook, PhD  
 COMMENT CABES Genome Facility  
 UC Davis, Plant Pathology  
 One Shields Ave, Davis, CA 95616, USA  
 Tel: 530 754 6561  
 Fax: 530 754 6617  
 Email: drcook@ucdavis.edu  
 Seq primer: ACGGTACCGACATATGCC.  
 Location/Qualifiers  
 1..753  
 /organism="Vitis vinifera"  
 /mol\_type="mRNA"  
 /cultivar="Cabernet Sauvignon"  
 /db\_xref="taxon:29760"  
 /clone="CAB10003\_11a\_Fa\_H06"  
 /sex="Hermaphrodite"  
 /dev\_stage="Pre-bloom"  
 /lab\_host="DH5alpha"  
 /clone\_lib="Cabernet Sauvignon Flower Pre-bloom - CAB1"  
 /note="Organ: Flower - Pre-bloom; Vector: pDMR; Site: 1;  
 SfiI; Site 2: SfiI; CAB1 is a cDNA library of Vitis  
 vinifera cv. 'Cabernet Sauvignon' Clone 8 flowers. Samples  
 were collected approximately eleven days before onset of  
 bloom (clusters at this stage were fully developed and  
 flowers with calypters or caps still attached. Sampled  
 vines were located at the University of California, Davis,  
 Experimental Vineyard. cDNAs were made by oligo-dT priming  
 and directionally cloned. 5' and 3' adaptors were used in  
 cloning as follows:  
 5'-AAGCAGTGATCAACGCCAGTGGCCATACGGCGGG-3' and  
 5'-ATTCTAGCGCGGCGGCGGACATG-3' (30)NN-3'. Library was  
 constructed using the Clontech Creator SMART kit and  
 size-selected to contain the 0.5-3 kb size fraction."

Query Match 28.2%; Score 451.4; DB 14; Length 753;  
 Best Local Similarity 77.2%; Pred. No. 4.4e-63;  
 Matches 548; Conservative 0; Mismatches 162; Indels 0; Gaps 0;

QY 714 TGTATTTCGCGCTGTTAGTCTAGCCTGAGGATATGAACCAACCAAGATTTGCTC 773  
 Db 1 TGTCTTTCGAGAGGCTTTTATGACGAGGAGGATTTGAGACAAAGAAATTTGGG 60  
 QY 774 AATCTGCACTTGAAGCTGCGCAGAGGCTCTGGATTTGTTGGAAATTTGGTGGG 833  
 Db 61 AAAACTGGACCTTAAACCTGCGCAGAGGCTCTAGATGTGGCTGTGGCATTTGGAGG 120  
 QY 834 TGACTTTTATACGCGGAGACCTTTGATGTGAGTGTGATTTGATTTGATCTCTCGTTAA 893  
 Db 121 TGACTTCTATATGCGAGAGACCTTTGATGTGAGTGTGATTTGATTTGATTTGATTT 180  
 QY 894 TATGATTTCTTTCGCTTGAAGCTTTATTTGGCTTAAATGCTGTGATTTGAGGT 953  
 Db 181 TATGTTTCTTTCGCTTGAAGCTTTGATTTGAGTGTGATTTGATTTGATTTGATTT 240  
 QY 954 AGCAGATTGCACCAAGATAAATACCTGCTGATTAATCTTTTGTATGTCATCTATAGCCGTGA 1013

Db 241 AGCTGACTGCACCAAGAAATCTATCTGCAACAACATTTGATGTGATCTACAGCGTGA 300  
 QY 1014 CACCAATTCGATATTTAGGACCAAGCTGCTGTTTGTAGATCTCTACAAATGGTTGAA 1073  
 Db 301 CACCAATTCGATATTTAGGACCAAGCTGCTGTTTGTAGATCTCTACAAATGGTTGAA 360  
 QY 1074 GCCAGAGGTAAGTTCTTAATCACTGATTAATCTGCAAGAAAGCTGTCACCTCCCTGAA 1133  
 Db 361 GCCAGAGGTAAGTTCTGATTTAGTATTAATCTGCAAGAAAGCTGTCACCTCCCTGAA 420  
 QY 1134 ATTCCGCCCTTACATTAAGCAGAGGGAATGATCTCCATGATTTGAAGAAATATGGGA 1193  
 Db 421 ATTTCGAGATATATTAACAAGAGGATGATGATCTTCAAGATTTGGAAGCATATGGTGA 480  
 QY 1194 GATGCTTAAAGATCTGATTTGTTGATGTTCTTCCGAGGATAGAACTAGCAGCTTCAT 1253  
 Db 481 GATGCTTAGAGATCTGCTTTCATTTGAGGTTCATTTGAGAGGATAGAACTAGCAGTCTCT 540  
 QY 1254 TCGAGTCTTACGGAAGGAACCTAGAGACTGTTGAGAGGAAAGGATGTTGATTTAGTGA 1313  
 Db 541 ACAAGTCTGCGAGAGGAGTTAAATGCTTTGAGACAAACAAGAAAGATTTGTTTCAGA 600  
 QY 1314 TTCTCTGAGAGGATTAACAATGACATTTGTTGAGGTTGGAATGATTAAGTTGCGGAGGAC 1373  
 Db 601 CTTCTCTGAAGACCACTACAAATGAAATAGTTGATGTTGGAAGGCTTAAGCTGATGAGGAG 660  
 QY 1374 TGCCAAAGGCTGAGCAAGCATGGGCTCTGTTGCTTGGCAAGAAAGATGAA 1423  
 Db 661 TTCTATCTGCTGAGCAGAGATGGGGCTTTTCATTTGCAAGAAAGGATGAA 710

RESULT 11  
 LOCUS CB970167  
 DEFINITION CB970167 753 bp mRNA linear EST 08-SEP-2003  
 Sweet Orange (Citrus sinensis (L.) Osbeck) Citrus sinensis cDNA  
 clone MVP-65\_G10 5', mRNA sequence.  
 ACCESSION CB970167  
 VERSION CB970167.1 GI:34523438  
 KEYWORDS EST.  
 SOURCE Citrus sinensis  
 ORGANISM Citrus sinensis  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosid II; Sapindales; Rutaceae; Citrus.  
 1 (bases 1 to 720)  
 Chaparro, J.X., Bausher, M.G., Dang, P., Hunter, W.B., McKenzie, C.L.,  
 Niedz, R.P., and Shatters, R.G. Jr.  
 Expressed Sequence Tags from Immature Ovaries of Valencia Sweet  
 Orange (Citrus sinensis (L.) Osbeck)  
 Unpublished (2003)  
 JOURNAL Contact: Chaparro, JX  
 COMMENT Horticulture and Breeding Research Unit  
 USDA, ARS, U. S. Horticultural Research Laboratory  
 2001, South Rock Road, Fort Pierce, FL 34945, USA.  
 Tel: 772 462 5830  
 Fax: 772 462 5986  
 Email: jchaparro@usbrl.ars.usda.gov  
 Seq primer: T3 Primer.  
 Location/Qualifiers  
 1..720  
 /organism="Citrus sinensis"  
 /mol\_type="mRNA"  
 /cultivar="Valencia"  
 /db\_xref="taxon:2711"  
 /clone="MVP-65\_G10"  
 /tissue\_type="Immature ovaries"  
 /clone\_lib="Immature Ovaries from field-collected Valencia  
 Sweet Orange (Citrus sinensis (L.) Osbeck)"  
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
 XhoI; Standard library construction protocols from  
 Stratagene cDNA synthesis kit (cat.# 200401-5) and Uni-ZAP



ORIGIN		Query Match		27.5%; Score 440.6; DB 14; Length 720;	
		Best Local Similarity		75.8%; Pred. No. 2.4e-61;	
		Matches 545; Conservative		0; Mismatches 174; Indels 0; Gaps 0;	
		XR vector kit (cat. # 237211) were followed using poly(A) RNA.			
QY		187	CTGGATTTCATTTGAGAGTGTCTATCAAGAAGATGAAGTAATCAATGGCCACTACAAAAT	246	
DB		2	CTGGACATTATTGACAGTGTCTATCAAGAAGATGAAGTAATCAATGGCCACTACAAAAT	61	
QY		247	GTCAAGTTTATGTCGATGATGATCTCTCCACTCTCAGTTTCCACCACTTCATTG	306	
DB		62	GTCAAAATTCATGTCGATGATGATCTCTCCACTCTCAGTTTCTCTGAGATTCAGTG	121	
QY		307	GATGTGATTTCTCCAAATGGTTACTCATGATATCTTTCTGATGAAGAGTGGAAAATTTG	366	
DB		122	GACATGATGTTCTCAAAATGGTTACTCATGATATCTTTCTGATGAAGAGTGGAAAATTTG	181	
QY		367	GTGGAAGAAATGTTGAATGGTTGAAGCCAGGGGTTACATTTCTTCAGAGAAATCTTGT	426	
DB		182	GCAGAAAGATGTTTAAATGGTTAAAGGTCGGTGGGTACATCTTCTCAGGAGTCATGT	241	
QY		427	TTCCATCAATCTGGGATCAACAGCAGAAAGCAATCCCACTCTACCGTCACTAGG	486	
DB		242	TTTCATCAATCTGGGATCAACAGCAGAAAGCAATCCCACTCTACCGTCACTAGG	301	
QY		487	TTCTACACTAAGGCTTCAAGAGTGTCTATTCAGAGATGGATCTGGAAATCTTTATGAG	546	
DB		302	TTTATTCAGAGTGTCTATTCAGAGATGGATCTTTTCAAGAGATGGATCTTTTATGAG	361	
QY		547	CTCTCCCTACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	606	
DB		362	CTATCACTGTTGTTTCAAGTGTCTATTCAGAGATGGATCTTTTCAAGAGATGGATCT	421	
QY		607	CAGATTAGTGTGTTGTCAGAAAGTTGATTTCTAAGAGTGAAGGGTTCAGCGATT	666	
DB		422	CAGATTAGTGTGTTGTCAGAAAGTTGATTTCTAAGAGTGAAGGGTTCAGCGATT	481	
QY		667	CTGGATCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT	726	
DB		482	CTGGATCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT	541	
QY		727	GGTATGTTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT	786	
DB		542	GGTATGTTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT	601	
QY		787	AGCCCTGGCCAGAGGTTCTGGAGTTGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	846	
DB		602	AGCCCTGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	661	
QY		847	GGCGAGACCTTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGAT	905	
DB		662	GCTGACAGTTTGTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGAT	720	
RESULT 12		BM302134		601 bp mRNA linear EST 22-JAN-2002	
LOCUS		MCA047F06		23452 Ice plant Lambda Uni-Zap XR expression library, 0	
DEFINITION				hours NaCl treatment prescreened for removal of highly abundant transcripts Mesembryanthemum crystallinum cDNA clone MCA047F06 5, mRNA sequence.	
ACCESSION		BM302134		GI:18024509	
VERSION		BM302134.1			
KEYWORDS		Mesembryanthemum crystallinum (common iceplant)			
SOURCE		Mesembryanthemum crystallinum			
ORGANISM		Eukaryote, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta,			
		Spermatophyta, Magnoliophyta, eudicotyledons, core eudicots,			
		Caryophyllales, Alismaceae, Mesembryanthemum.			
REFERENCE		1 (bases 1 to 601)			

Cushman, J.C.		An expressed sequence tag database for the common ice plant,	
Mesembryanthemum crystallinum			
Unpublished (1997)			
Contact: Cushman JC			
Department of Biochemistry			
University of Nevada			
MS200, Reno, NV 89557-0014, USA			
Tel: 775-784-1918			
Fax: 775-784-1650			
Email: jcushman@unr.edu			
PCR Primers			
FORWARD: T3 20mer			
BACKWARD: T7 21mer			
Plate: 047 row: F column: 06			
Seq primer: T3 20mer			
High quality sequence stop: 601.			
Location/Qualifiers			
1..601			
/organism="Mesembryanthemum crystallinum"			
/mol_type="mRNA"			
/db_xref="taxon:3544"			
/clone="MCA047F06"			
/tissue_type="leaf"			
/dev_stage="six-week-old"			
/clone_lib="Ice plant Lambda Uni-Zap XR expression			
library, 0 hours NaCl treatment prescreened for removal of			
highly abundant transcripts"			
/notes="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site_1:			
EcoRI; Site_2: XhoI"			
ORIGIN			
Query Match		27.2%; Score 435.2; DB 12; Length 601;	
Best Local Similarity		82.7%; Pred. No. 1.9e-60;	
Matches 497; Conservative		0; Mismatches 104; Indels 0; Gaps 0;	
QY	192	TTTCATTGAGAGTCTATCAAGAAGATCAAGTAATCAATGGCCACTACAAAATGTCAA	251
DB	1	TTTCATTGAGAGTCTATTAAGAAGAAATCAAAAGCATAAATGGTCAATTACGAAAATGTCAA	60
QY	252	GTTCATGTCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	311
DB	61	GTTCATGTCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	120
QY	312	GATATTTCCAAATGGTTTACTCATGATATCTTTCTGATGAAGAGTGGAAAATTTGGTTGA	371
DB	121	GATATTTCCAAATGGTTGTCATGATATCTTTCTGACGATGAGGTTGAGACTCTAGTTAA	180
QY	372	AGCAATGTTGAAATGGTTTGAAGCCAGGGGTTACATTTCTTCAGAGAAATCTTGTTTCCA	431
DB	181	AGATATGTTGAAATGGTTTGAAGCCAGGGGCTATATTTCTTCAGAGAGATCCTGTTTCCA	240
QY	432	TCATTTCTGGGATCAAAACGCAAAAGCATCCCACTACCGTGAACCTAGGTTCTA	491
DB	241	TCATTTCTGGGATCAAAACGCAAAAGCATCCCACTACCGTGAACCTAGGTTTCTA	300
QY	492	CATTAAGGCTTCAAGAGTGTCAATTGCAAGATGGAATCTGGAACTCTTATGAGCTCTC	551
DB	301	CATTAAGGCTTCAAGAGATGCCATATGCAAGATAATTCAGGAACTCTATGAGCTTTC	360
QY	552	CCTACTTACTGCAATGTTATTTGAGCTTATGAGTATGAGTATGAGTATGAGTATGAGT	611
DB	361	CCTGGTTGGTTGTAATTCATTTGGAGCTTATGTTAGAAAACAAGAAATCAAAACCAAGT	420
QY	612	TAGTTGTTGTCGCAAAAGTTGATTTCTTAAGGATGATAAGGGGTTTCCAGCAATTTCTGGA	671
DB	421	TTGCTGTTATGCGAGAAAGTTTGACTCAGAAGATGATAAAGGTTTTCAGCGAATTTTTCGA	480
QY	672	TACTAGCAGTACAGTGTAAATAGCATTTCTGCATATGAGCGTGTATTTGGCCCTGGTTA	731
DB	481	TACTATCCAATATAATGATGTGGCAATTTTGGCCATCGAGCTGTCTTTGGCCATGTTTA	540
QY	732	TGTTAGCACTGGAGGATATGAACCAACAAGAGGTTTGTGTCAATGCTGGACTTGAAGCC	791



```

Db      541 TGTGACACAGAGGAGTTCACAAACAAAGAGATTGTGTCGAGCTTGACCTGACGCC 600
Qy      792 T 792
Db      601 T 601

RESULT 13
LOCUS   BE040284
DEFINITION OEO3A05 OE Oryza sativa cDNA 5' similar to
          s-adenosyl-methionine-sterol-c-methyltransferase, mRNA sequence.
ACCESSION BE040284
VERSION   BE040284.1 GI:8335300
KEYWORDS EST.
SOURCE   Oryza sativa
ORGANISM Oryza sativa
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
          Eriacoidae; Oryzaceae; Oryza.
          1 (bases 1 to 822)
          Bohnert, H.J., Borchert, C., Brazille, S., Brooks, J., Eaton, M.,
          Ferreira, H., Kawasaki, S., McCollough, A., Michalowski, C.B.,
          Palacio, C., Scara, G., Wheeler, M. and Zepeda, G.R.
          Functional Genomics of Plant Stress Tolerance
          Unpublished (2000)
          Contact: Michalowski, C.B.
          University of Arizona
          Bio Sciences West room 513, Tucson, AZ 85721, USA
          Tel: 520-621-7982
          Fax: 520-621-1697
          Email: cbm@u.arizona.edu
          An open reading frame exists.
          Insert Length: 1 Std Error: 0.00.
          Location/Qualifiers
            1..822
              /organism="Oryza sativa"
              /mol_type="mRNA"
              /strain="Porkali"
              /db_xref="taxon:4530"
              /tissue_type="roots"
              /dev_stage="1 week"
              /clone_lib="OR"
              /note="2-3 d 150mM NaCl"

ORIGIN
Query Match      26.88; Score 429; DB 10; Length 822;
Best Local Similarity 71.74; Pred. NO. 1.7e-59;
Matches 589; Conservative 0; Mismatches 230; Indels 2; Gaps 2;

Qy      475 CGTGAACCTAGTTCTACACTAAGGCTTCAAGAGTGTCAATTTGCAAGATGGATCTGGA 534
Db      1 CGGAGGCAAGGTTTGAGACTAAGTGTGTTAAGAGTGTCAAGCTTTGATCAAGATGGG 60
Qy      535 AACTCTTATGAGCTCTCCCTACTTACTGCAATGTATTGGGCTTATGTGCAAGAACAG 594
Db      61 AATTCCTTGAACTCTCTACTTACTTCTCAAGTGTGTGAGCTTACTGTGAAGCAAG 120
Qy      595 AAAACACGACCAAGATTTAGTTGGTTGTGCAAAAAGTTGATTTCAAGGATGATAAGGGG 654
Db      121 AAAAATCAAAACCAAGATATGTGGCTATGSCAAAAGTTGATTCAACAGAGATCGGGG 180
Qy      655 TTCAGGAGTTCTGAGTACTAGCCAGTCAAGTGTATAGATCTCGGTATATGCGGT 714
Db      181 TTTCAAGATTTTGGCAATGTGCAATCAAGCAGTGGATATACGCTATGAACGC 240
Qy      715 GTATTGGCCCTGGTTAGTTAGTCACTGAGGATATGAACCAACAAAGATTTGTCTCA 774
Db      241 ATCTTGGAGAGGCTTGTGAGCACTGTGTGCAATTTGAACTACAAAGATTTGTGAC-300
Qy      775 ATGCTGCACTTGAAGCTCGCCAGAGGTCCTCGATGTGTGGATTTGGGATTTGGTGAAGT 834

```

```

Db      301 AGGCTGGATCTCAAACTGCGCAGAACGTTCTTGATGTTGGATGTGGAATTGGGGCGGT 360
Qy      835 GACTTTACATGCGGAGACCTTTGATGTTGAGTTGTTGAGTTGTTGATCTCTCCCTTAAT 894
Db      361 GATTTTATATGCGTGCACAAAGTATGATGTTCAATGTTGTTGTTGTTGATCTCTCGATAAC 420
Qy      895 ATGATTTCTTTGCGCTTGAGCGTTCTTATTTGGCTTTAAATGTTGCTGTTGAGTTTGAAGTA 954
Db      421 ATGGTTTCTTTTGCACCTTGAGCGTGTCTATTGGCGTAAAGTCTCAGTTGAGTTTGAAGTC 480
Qy      955 GCAGATTCACCAAGATAAACTACCTCGATAAATCTTTTGTGATGTCATCTATAGCCGTGAC 1014
Db      481 GCTGATTCGACCAAAAGACATACCCAGACACACGTTTGACGTATCTATACAGTCTGTGAT 540
Qy      1015 ACCATTCTGCATATTCAGGACAGCGCTGTTGTTGATGTCCTTCTACAAATGTTGAAG 1074
Db      541 ACTATCTCTACATACAGATAAAGCTTCACTATTTAAAGTTTCTTCAAGTGGCTCAA 600
Qy      1075 CCAGAGGTAAGGTTCTTAATCAGTGATTACTGCAAGAAAGCTGGTCCACCTCACCTGAA 1134
Db      601 CTGGGGTGAAGTCTTAATTTAGTATTAATCTGCAAGTGGCTTGGAAACCTTCAAGAG 660
Qy      1135 TTGCGCGCTTACATTAAGCAGAGCGGATATGATCTCCATGATGTAAAGGAATATGGGCG 1194
Db      661 TTGCGAGCTTACATTAAGCAAGGGGTTATGACCTTAAACAGCTCAGGGCTTACGGCAG 720
Qy      1195 ATGCTTAAAGATGCTGGAATTTGTTGATGTTCTTGGCGGATAGAACTGAGCAGTTCAAT 1254
Db      721 ATGCTTCAA-ATGCTGTTTCCATTCATGTCATTTGTTGAAGCCGACCGATCAGTTC-CT 778
Qy      1255 CGAGTTCTAGGAGGAACTAGAGACTGTTGAGAGGAAA 1295
Db      779 CGAGTTCTAGAGGAGGAGCTGTGTAAGATTGAAGACCAAA 819

RESULT 14
LOCUS   BQ120291
DEFINITION EST605867 mixed potato tissues Solanum tuberosum cDNA clone STMEP82
          5' end, mRNA sequence.
ACCESSION BQ120291
VERSION   BQ120291.2 GI:21919928
KEYWORDS EST.
SOURCE   Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          asterids; lamids; Solanales; Solanaceae; Solanum.
          Buell, C.R., Hart, A., Baker, B., Tankale, S., Fry, W., Smart, C.,
          Restrepo, S., Griffiths, H., van der Hoeven, R., Tsai, J. and
          Karamycheva, S.A.
          Generation of a set of potato cDNA clones for microarray analyses
          Unpublished (2002)
          On Apr 17, 2002 this sequence version replaced gi:20172253.
          Contact: Robin Buell
          The Institute for Genomic Research
          9712 Medical Center Dr, Rockville, MD 20850, USA
          Email: potato-array@igr.org
          This clone can be obtained from the University of Arizona Genomics
          Institute. Orders can be made through URL:
          http://genome.arizona.edu/orders/
          Seq primer: 13.
          Location/Qualifiers
            1..706
              /organism="Solanum tuberosum"
              /mol_type="mRNA"
              /cultivat="Kenuebac or Binjete"
              /db_xref="taxon:4113"
              /clone="STMEP82"
              /tissue_type="mixed tissues"
              /lab_host="SOLR"
              /clone_lib="mixed potato tissues"

```

ORIGIN

Query Match 26.7%; Score 427.2; DB 13; Length 706;  
 Best Local Similarity 75.4%; Pred. No. 3.5e-59;  
 Matches 531; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

464 CCACCACTACCGTGAACCTAGGTTCTACATTAAGCCCTTCAAGAGTGTCATTGCAAG 523  
 Db 2 CAACCCATTATCGGAGCTAGATTCTTACACAAAGGTGTTAAAGAAATGTCATATAATG 61

524 ATGATCTGGAATCTTATAGCTCTCCCTACTAGCTGCAAAATGATTCGACCTATG 583  
 Db 62 CTGTGTGATGTAATCATTTTGAATCTTCTCATGTGTCAGTGCAATGAGCTTATG 121

584 TCAGAAACAAAGAAACCAAGACAGATTAGTGTGTTGTCGCAAAAGTTGATCTAAG 643  
 Db 122 TTAAGAAACAAAGAAATCAGATCAGATTGTTGGCTATGTCGCAAAAGTTAATCTGAG 181

644 ATGATGAGGGGTTCACGAGATTCTGTAATAGCCAGTACAGTGAATGATGCTTCG 703  
 Db 182 ATGACAGGGGATTCACGGTTCTTGCAAAATGTTCAATACAAATGATGCTGCTATG 241

704 GATATGACCGTGTATTTGGCCCTGTTATGATGACCTGAGGATGATGAAACCAAG 763  
 Db 242 GATATGACCGTGTATTTGGGAGAGTTATGATGACGACGAGGACTTATGATCACAAG 301

764 AGTTTGTCTCAATCTGGAATCTGAAAGCTGCGCAGAGGTCTGATGTTGTTGGA 823  
 Db 302 AATGCTTCTATGTTGATCTTCAACCTGCGCAAAAGTCTTATGATGCTGCGGA 361

824 TTGGTGAAGGTGATTTTATGATGCGGAGACCTTATGATGAGGTTGATGATGATC 883  
 Db 362 TAGTGTGAGGAGTCTTATGATGCTGAGATTTATGATGATGATGATGATGATGAT 421

884 TCTCGTTTATATGATTTCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT 943  
 Db 422 TCTCGTTTATATGATTTCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT 481

944 AGTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1003  
 Db 482 AATTTGAGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 541

1004 ATAGCGGTGACACCATTCGATATTCAGGACAGCTGCTGCTGCTGCTGCTGCTGCT 1063  
 Db 542 ACAGTGAAGACATATTCCTTACATCCAGACAAACCCAGCTTATTCAGATCTTCTACA 601

1064 AATGTTGAAGCCAGGAGTAAAGTCTTAAATCAGTGAATTCGCAAGAAAGCTGCTCAC 1123  
 Db 602 GGTGCTGAAACAGGAGGCAAGTCTCTCAATAGTGAATTAAGTGAATTAAGTGAAT 661

1124 CCTCAGCTGAATTCGCGCTTACATTAAGCAGAGGGGATGAT 1167  
 Db 662 CATCGACAGATTTTCTGATGATTAAGCAAGGGGTTATGAT 705

RESULT 15  
 BQ965516/c  
 LOCUS  
 DEFINITION  
 BQ965516  
 BQ965516.1 GI:22382621  
 EST.  
 Helianthus annuus (common sunflower)  
 ORGANISM  
 Helianthus annuus  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; campanulids; Asterales; Asteraceae; Asteroidae;

Heliantheae; Helianthus.  
 1 (bases 1 to 733)  
 Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L.,  
 Lin, H., van Dommelen, M., Lavelle, D., Chevalier, P., Ziegler, J.,  
 Ellison, P., Kolman, J., Slabaugh, M.S., Livingston, K., Zhou, Y.,  
 Lai, Z., Church, S., Jackson, L. and Bradford, K.  
 Lettuce and Sunflower ESTs from the Composite Genome Project  
 http://compgenome.ucdavis.edu/  
 Unpublished (2002)  
 Contact: Alexander Kozik [R.W.Michelmore]  
 Department of Vegetable Crops, R.W.Michelmore Lab  
 University of California at Davis (UCD)  
 Amundson Hall, UCD, Davis, CA 95616, USA  
 Tel: 1-(530)-742-1742  
 Fax: 1-(530)-752-9659  
 Email: akozik@ucdavis.edu [michelmore@vegmail.ucdavis.edu]  
 belongs to contig QH\_CA\_Contig1096, see http://cgpdb.ucdavis.edu/  
 for details  
 Plate: QHB22 row: C column: 02.

FEATURES  
 source  
 1..733  
 /location/Qualifiers  
 /organism="Helianthus annuus"  
 /mol\_type="mRNA"  
 /cultivar="RHA801"  
 /db\_xref="taxon:4232"  
 /clone="QH22C02"  
 /lab\_host="E.coli"  
 /clone.lib="OH ABCDI sunflower RHA801"  
 /notes="Vector: pBRCNDA51AB; The library was constructed  
 from 11 different sources of RNA from a single genotype.  
 Separate cDNAs were generated using primers that  
 incorporated unique 5' and 3' tags to distinguish each  
 source of RNA. cDNAs were then pooled, size-fractionated,  
 directionally cloned into a custom medium-copy vector and  
 transformations made with four size classes to minimize  
 size bias. Details of each source of RNA and library  
 construction can be obtained at http://cgpdb.ucdavis.edu/  
 TAG TISSUE=hulls  
 TAG LIB=OH ABCDI sunflower RHA801  
 TAG\_SEQ=GGTGTGCGG"

ORIGIN  
 Query Match 26.6%; Score 426; DB 13; Length 733;  
 Best Local Similarity 74.7%; Pred. No. 5.4e-59;  
 Matches 547; Conservative 0; Mismatches 182; Indels 3; Gaps 1;

Qy 249 CAAGTTTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 308  
 Db 733 CGAATTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 674

Qy 309 TGTGATATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 368  
 Db 673 TTGATGATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 614

Qy 369 TGAAGAAATGTTGAAGTGTGAAGCCAGGGGGTTACATTTCTTCAGAGAAATCTGTTT 428  
 Db 613 TGAAGAAATTTTGAAGTGTGAAGTGTGAAGTGTGAAGTGTGAAGTGTGAAGTGT 554

Qy 429 CCATCAATCTGGGATCACAACCAAGCAATCCACCCACTCCGCTGAGACCTAGGTT 488  
 Db 553 TCACGCTGTGAGACCAATAGCAAGCAATCCACACATCTCCGCTGAGACCTAGGTT 494

Qy 489 CTACACTAAGGCTTCAAGAGTGTCAATTTGCAAGATGATGATGATGATGATGATGAT 548  
 Db 493 CTACACCAAGGCTTTTAAAGAAATGTCACCACTGATGATGATGATGATGATGATGAT 434

Qy 549 CTCCTCTACTTACTGCAATATGATTTGGAGCTTATGTCAGAAACACAGAAACACAGAACCA 608  
 Db 433 TAGGCTTAACCTGATCGAATGCAATTCGAGCTTATGTCAGCAACCAAGAAATCAAAATCA 374

Qy 609 GATTAGTGTGCTGCAAGAGTGTGCTTAAAGTGTGCTTAAAGTGTGCTTAAAGTGT 665  
 Db 373 GATCTCTGTGATATGCGAGAGGTTGCGGCAATTCGAGACCGATAGGACTTCCAGGAT 314

```
Qy 666 TCTGGATCTAGCCAGTACAAAGTGTAAATAGCAATCTCGATATGAGCGTGTATTTGGCCC 725
Db |||||
313 CTTGGATAACGTACAGTACAAAGCAATGGCATCTGGCTAGCAACGCATATTTGGACC 254
Qy 726 TGGTTATGTTAGCACTGGAGGATATGAACCAACCAAGAGTTTGTCTCAATGCTGGACTT 785
Db |||||
253 AGGGTTTGTGACCAAGGAGGATAGACACGACAAAGAAATTTTGGCAATGCTGGATCT 194
Qy 786 GAAGCTTGGCCAGAGGCTCTGGATGTTGGTGTGGAAATGGTGGAGGTGACTTTTACAT 845
Db |||||
193 GAAACCGGGGCAGAAAGTTTAGACGTTGGATGTGGAATCGGGGAGGCGCACTTCTATAT 134
Qy 846 GCGGAGACCTTTGATGTTGAGGTTGTGGATTTGATCTCTCGTTAAATATGATTTTCCTT 905
Db |||||
133 GCGTGAAGAACTTTGACGTTGATGTTTGGCAATGATCTTCTGCAATATGATCTCTTT 74
Qy 906 TGCCCTTGAGCGTTCTATTGGGCTTAAATGTGCTGTTGAGTTGAGGTAGCAGATTGCAC 965
Db |||||
73 TGCTCTTGACGTGCATTTGGGCTCCAAATGTTCAATTTGAATTCGAGTTGCTGACTGTAC 14
Qy 966 CAAGATAAATA 977
Db |||||
13 CAAGAAATCATA 2
```

Search completed: August 2, 2004, 08:17:45  
Job time : 4447 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2004, 04:36:43 ; Search time 6481 Seconds  
(without alignments)  
10713.708 Million cell updates/sec

Title: US-10-031-331B-39

Perfect score: 1602

Sequence: 1 cacaccgtgtatttaaccat.....aaaaaaaaaaaaaaaaaaaa 1602

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_hg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_ma.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vi.\*

30: em\_hgt\_hum.\*

31: em\_hgt\_inv.\*

32: em\_hgt\_other.\*

33: em\_hgt\_mus.\*

34: em\_hgt\_pln.\*

35: em\_hgt\_rtd.\*

36: em\_hgt\_mam.\*

37: em\_hgt\_vrt.\*

38: em\_sy.\*

39: em\_hgt\_hum.\*

40: em\_hgt\_mus.\*

41: em\_hgt\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1602	100.0	1602	6	BD084055
2	1602	100.0	1602	6	BD084055
3	1602	100.0	1779	8	AB080186
4	1138.6	71.1	2235	6	AX250551
5	1138.6	71.1	2252	8	AF237633
6	875	54.6	1794	8	AF328858
7	857.4	53.5	1870	8	AB090883
8	853.4	53.2	1656	8	AF428454
9	851.8	53.2	1459	8	AY133811
10	851.8	53.2	1660	8	AY063866
11	834	52.1	1770	8	AY058175
12	834	52.1	1809	8	AF367299
13	834	52.1	1878	8	AF197940
14	832.2	52.0	1476	8	AY091683
15	832.2	52.0	1911	8	AY093093
16	822.2	51.3	1750	8	AY139479
17	766.8	47.9	2069	8	AK069137
18	748.4	46.7	1834	8	AY136372
19	700.6	43.7	1782	8	AY065971
20	621.2	38.8	1120	6	AX250553
21	302.6	18.9	724	8	HVU234432
22	209	13.0	3221	8	AK102037
23	189.4	11.8	74968	8	AB019230
24	187.2	11.7	29508	8	AC079676
25	187.2	11.7	105937	8	AC012679
26	186.6	11.6	95295	8	AC030889
27	180.6	11.3	16886	8	AP003248
28	168.2	10.5	144305	2	AC137078
29	166.6	10.4	141715	2	AC130725
30	166.6	10.4	158911	2	AC136492
31	160.2	10.0	111754	2	AC138199
32	149.4	9.3	280	6	AR251394
33	149.4	9.3	286	6	AR252133
34	125.8	7.9	274	6	AR248084
35	109.2	6.8	260	6	AR245517
36	101	6.3	174	6	AR250330
37	100.4	6.3	274	6	AR250330
38	75.8	4.7	110000	2	PFMAL13_16
39	74.4	4.6	1220	3	AY429590
40	63.2	3.9	23614	3	U64834
41	63.2	3.9	140702	2	AC006888
42	58.4	3.6	2522	8	ATU80192
43	57	3.6	171	6	AR248566
44	56.6	3.5	3016	3	DDU11631
45	56.6	3.5	6424	3	AC116924

# ALIGNMENTS

RESULT 1	BD084055	1602 bp	DNA
LOCUS	BD084055	Environmental stress-tolerant gene.	linear
DEFINITION	BD084055	Environmental stress-tolerant gene.	
ACCESSION	BD084055.1	GI:22629665	
VERSION	JP 2001333784-A/20.		
KEYWORDS	unidentified		
SOURCE	unclassified		
ORGANISM	unclassified		
REFERENCE	1 (bases 1 to 1602)		
AUTHORS	Yamada, T., Ohtsuki, Y., and Saito, T.		
TITLE	Environmental stress-tolerant gene		
JOURNAL	Patent: JP 2001333784-A 20 04-DEC-2001;		
	JAPAN SCIENCE AND TECHNOLOGY CORP		

COMMENT		OS		Sueada japonica	
PN	JP 2001333784-A/20	PN	JP 2001333784-A/20	PN	JP 2001333784-A/20
PD	04-DEC-2001	PD	04-DEC-2001	PD	04-DEC-2001
PF	19-JUL-2000 JP 2000219649	PF	19-JUL-2000 JP 2000219649	PF	19-JUL-2000 JP 2000219649
PI	AKIYO YAMADA, YOSHIHIRO OZEKI, TAKEO SAITO	PI	AKIYO YAMADA, YOSHIHIRO OZEKI, TAKEO SAITO	PI	AKIYO YAMADA, YOSHIHIRO OZEKI, TAKEO SAITO
PC	C12N15/09, A01H5/00, C07K14/415, C07K16/16, C12N5/10, C12P21/02//	PC	C12N15/09, A01H5/00, C07K14/415, C07K16/16, C12N5/10, C12P21/02//	PC	C12N15/09, A01H5/00, C07K14/415, C07K16/16, C12N5/10, C12P21/02//
PC	C12N1/21,	PC	C12N1/21,	PC	C12N1/21,
PC	C12P21/08, C12N15/00, C12N5/00	PC	C12P21/08, C12N15/00, C12N5/00	PC	C12P21/08, C12N15/00, C12N5/00
CC	Environmental stress-tolerant gene	CC	Environmental stress-tolerant gene	CC	Environmental stress-tolerant gene
PH	Location/Qualifiers	PH	Location/Qualifiers	PH	Location/Qualifiers
FT	Key (1) . (1419).	FT	Key (1) . (1419).	FT	Key (1) . (1419).
FEATURES		source		Location/Qualifiers	
				1. 1602	
				/organism="unidentified"	
				/mol type="genomic DNA"	
				/db_xref="taxon:32644"	
ORIGIN		Query Match		100.0%; Score 1602; DB 6; Length 1602;	
		Best Local Similarity		100.0%; Pred. No. 0;	
		Matches 1602; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1	CACACGTTGATTAAACCAATTGAAGCTATGATGCTCGAATTCCTCAAGCTTCTGATCTTGAC	60		
Db	1	CACACGTTGATTAAACCAATTGAAGCTATGATGCTCGAATTCCTCAAGCTTCTGATCTTGAC	60		
Qy	61	AAGAGAAAGCTCTGAGATCTTTCAATGCTTCGGCTCTTGAAGCAAAATGCTCTTG	120		
Db	61	AAGAGAAAGCTCTGAGATCTTTCAATGCTTCGGCTCTTGAAGCAAAATGCTCTTG	120		
Qy	121	GAACTTGGGGCTGATTTGGTCTGTTTACTGGTGAATTTGGCTGAGAAAGCTGGCAGGTT	180		
Db	121	GAACTTGGGGCTGATTTGGTCTGTTTACTGGTGAATTTGGCTGAGAAAGCTGGCAGGTT	180		
Qy	181	ATTGCTCTGATTTTCAATGAGGCTCTATCAAGAGAAATCAAGTAAATCAATGGCACTAC	240		
Db	181	ATTGCTCTGATTTTCAATGAGGCTCTATCAAGAGAAATCAAGTAAATCAATGGCACTAC	240		
Qy	241	AAAATGTCAAGTTATGTTGCTGATGATCTTCCCACTCTCAGTTTCCCAACCAT	300		
Db	241	AAAATGTCAAGTTATGTTGCTGATGATCTTCCCACTCTCAGTTTCCCAACCAT	300		
Qy	301	TCAATGATGATATCTTCCAAATGGTTACTCATGATCTTCTCATGATGAGAGGTTGAA	360		
Db	301	TCAATGATGATATCTTCCAAATGGTTACTCATGATCTTCTCATGATGAGAGGTTGAA	360		
Qy	361	AAATTTGGTGAAGAAATGTTGAATGGTTGAAGCAAGGGGTTACATTTCTTCAGAA	420		
Db	361	AAATTTGGTGAAGAAATGTTGAATGGTTGAAGCAAGGGGTTACATTTCTTCAGAA	420		
Qy	421	TCTTGTTCATCAATCTGGGATCAAAACGCAAAAGCAATCCCACTACCGTGAA	480		
Db	421	TCTTGTTCATCAATCTGGGATCAAAACGCAAAAGCAATCCCACTACCGTGAA	480		
Qy	481	CCTAGGTTCTACACTAAGGCTTCAAGAGTGTCAATTTCAAGATGATCTGAAACTCT	540		
Db	481	CCTAGGTTCTACACTAAGGCTTCAAGAGTGTCAATTTCAAGATGATCTGAAACTCT	540		
Qy	541	TATGAGCTCTCCCTACTAGCTGCAAAATGTTATGGAGCTTATGTTCAGAAACAAAGAAC	600		
Db	541	TATGAGCTCTCCCTACTAGCTGCAAAATGTTATGGAGCTTATGTTCAGAAACAAAGAAC	600		
Qy	601	CAGAACCAAGATTAGTTGGTTGGCAAAAGTTGATTTCAAGGATGATTAAGGGGTTCCAG	660		
Db	601	CAGAACCAAGATTAGTTGGTTGGCAAAAGTTGATTTCAAGGATGATTAAGGGGTTCCAG	660		
Qy	661	CGATTTCTGATCTAGCTAGCTCAAGTGTATAGCATCTTGGATATGAGCGGTATTT	720		
Db	661	CGATTTCTGATCTAGCTAGCTCAAGTGTATAGCATCTTGGATATGAGCGGTATTT	720		
Qy	721	GGCCCTGTTATGTTAGCTAGCTAGGATGAAACCAACCAAGAGTTTGTCTCAATCTG	780		
Db	721	GGCCCTGTTATGTTAGCTAGCTAGGATGAAACCAACCAAGAGTTTGTCTCAATCTG	780		

Qy	781	GACTTGAAGCTTGGCCAGAAAGTCTTGGATGTTGGTGTGGAAATTTGGTGGAGTGTGACTTT	840
Db	781	GACTTGAAGCTTGGCCAGAAAGTCTTGGATGTTGGTGTGGAAATTTGGTGGAGTGTGACTTT	840
Qy	841	TACATGGCGGAGACCTTTGATGTTGAGGTTGTTGGAATTTGATCTCTCCGTTAAATATGATT	900
Db	841	TACATGGCGGAGACCTTTGATGTTGAGGTTGTTGGAATTTGATCTCTCCGTTAAATATGATT	900
Qy	901	TCCTTGGCCCTTGAAGCTTCTATGCGCTTAATGCTGTTGAGTTTGAAGTGTGAGTACAGAT	960
Db	901	TCCTTGGCCCTTGAAGCTTCTATGCGCTTAATGCTGTTGAGTTTGAAGTGTGAGTACAGAT	960
Qy	961	TGCACCAAGATAAACTACCTGATAAATCTTTTGAATGTCATCTATAGCGGTGACACCAT	1020
Db	961	TGCACCAAGATAAACTACCTGATAAATCTTTTGAATGTCATCTATAGCGGTGACACCAT	1020
Qy	1021	CTGCATATTCAGGACAAAGCTGGTGTGTTTGAATGCTTCTACAAATGGTTGAAGCCAGCA	1080
Db	1021	CTGCATATTCAGGACAAAGCTGGTGTGTTTGAATGCTTCTACAAATGGTTGAAGCCAGCA	1080
Qy	1081	GSTAAAGTCTTAATCAGTGTACTGCAAGAAAGCTGGTCCACCTCACCCTCACTGAATTCGCC	1140
Db	1081	GSTAAAGTCTTAATCAGTGTACTGCAAGAAAGCTGGTCCACCTCACCCTCACTGAATTCGCC	1140
Qy	1141	GCTTACATTAAGCAGAGGGGATATGATCTCCATGATGTAAGCAATATATGGCAGATGCTT	1200
Db	1141	GCTTACATTAAGCAGAGGGGATATGATCTCCATGATGTAAGCAATATATGGCAGATGCTT	1200
Qy	1201	AAAGATGCTGATTTGTTGATGTTCTTCCGAGGATAGAACTGAGCAGTTCATTGAGTT	1260
Db	1201	AAAGATGCTGATTTGTTGATGTTCTTCCGAGGATAGAACTGAGCAGTTCATTGAGTT	1260
Qy	1261	CTACGAGGAACTAGAGACTGTTGAGAGAAAGATGTTGTTCAATTAAGTGTGATTTCTCT	1320
Db	1261	CTACGAGGAACTAGAGACTGTTGAGAGAAAGATGTTGTTCAATTAAGTGTGATTTCTCT	1320
Qy	1321	GAGGAGGATTACAAATGACATTTGTTGAGGTTGGAATGATAAGTTGGGAGGATCTGCCAAG	1380
Db	1321	GAGGAGGATTACAAATGACATTTGTTGAGGTTGGAATGATAAGTTGGGAGGATCTGCCAAG	1380
Qy	1381	GGTGAGCAACGATGGGCTCTGTTGTTCCGAGAGAGGATGAAAGTTCAGTTGCCGACT	1440
Db	1381	GGTGAGCAACGATGGGCTCTGTTGTTCCGAGAGAGGATGAAAGTTCAGTTGCCGACT	1440
Qy	1441	GGCAGCTGCTGATTTCTAGTATTAATCTTCAATGTTTTCATGTAATGATCTTCTACATGT	1500
Db	1441	GGCAGCTGCTGATTTCTAGTATTAATCTTCAATGTTTTCATGTAATGATCTTCTACATGT	1500
Qy	1501	AAATTTGCCAATTAAGTTCGATTCGACATCTGTAAGATGATTAATCATATTTTATCTTTT	1560
Db	1501	AAATTTGCCAATTAAGTTCGATTCGACATCTGTAAGATGATTAATCATATTTTATCTTTT	1560
Qy	1561	AATTAATCATGATTTATGCAAAAAA	1602
Db	1561	AATTAATCATGATTTATGCAAAAAA	1602

RESULT 2	BD093370	1602 bp	DNA	linear	PAT 27-AUG-2002
LOCUS	BD093370	Environmental stress-tolerant gene.			
DEFINITION	BD093370	Accession			
VERSION	BD093370.1	GI:22638958			
KEYWORDS	WO 0106006-A/20.				
SOURCE	unidentified				
ORGANISM	unclassified				
REFERENCE	1 (base 1 to 1602)				
AUTHOR	Yamada, A., Ozeki, Y. and Saito, T.				
TITLE	Environmental stress-tolerant gene				
JOURNAL	Patent: WO 0106006-A 20 25-JAN-2001;				
	JAPAN SCIENCE AND TECHNOLOGY CORP, AKIYO YAMADA, YOSHIHIRO OZEKI,				

COMMENT	TAKEO SAITO OS Sueada japonica PN WO 0106006-A/20 PD 25-JAN-2001 PR 19-JUL-2000 WO 2000JP004862 AKIYO YAMADA, YOSHIHIRO OZeki, TAKEO SAITO PC C1201/68, C12N15/29, C07K14/415, C07K16/16, C12P21/02, A01H5/00 CC FH Key Location/Qualifiers FT CDS (1). (1419).									
FEATURES	Location/Qualifiers 1..1602 /organism="unidentified" /mol_type="genomic DNA" /db_xref="taxon:32644"									
SOURCE	Query Match 100.0%; Score 1602; DB 6; Length 1602; Best Local Similarity 100.0%; Pred. No. 0; Matches 1602; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
ORIGIN	1 CACACGGTTGATTTAAACCAATGAAGCTATGATGCTCCGATCTCAAGCTTCTGATCTTGAC 60 1 CACACGGTTGATTTAAACCAATGAAGCTATGATGCTCCGATCTCAAGCTTCTGATCTTGAC 60 61 AAAGAAGAACGCTCTGAGATCTTCAATGCTCCGCTCTTGAAGGAAATGCCCTCTTG 120 61 AAAGAAGAACGCTCTGAGATCTTCAATGCTCCGCTCTTGAAGGAAATGCCCTCTTG 120 121 GAACTTGGGGCTGATTTGGTGGTTTAACTGGTGAATGGCTGAGAAAGCTGGCCAGGTT 180 121 GAACTTGGGGCTGATTTGGTGGTTTAACTGGTGAATGGCTGAGAAAGCTGGCCAGGTT 180 181 ATTGCTCTGGATTTCAATGAGAGTGCTATCAAGAAGAAATGAATCAATCAATGGCACTAC 240 181 ATTGCTCTGGATTTCAATGAGAGTGCTATCAAGAAGAAATGAATCAATCAATGGCACTAC 240 241 AAAATGTCMAAGTTATGTCGTGATGTCGATCTTCCCACTCTCAATTTCCCAACAT 300 241 AAAATGTCMAAGTTATGTCGTGATGTCGATCTTCCCACTCTCAATTTCCCAACAT 300 301 TCATTTGGATGATATTTCCCAATTTGGTGTACTCATGTATCTTCTCATGATCAAGAGGTGGA 360 301 TCATTTGGATGATATTTCCCAATTTGGTGTACTCATGTATCTTCTCATGATCAAGAGGTGGA 360 361 AATTGGTTGAAAGAAATGTTGAATGTTCAAGCCAGGGGTTTACATTTCTTCAGAGAA 420 361 AATTGGTTGAAAGAAATGTTGAATGTTCAAGCCAGGGGTTTACATTTCTTCAGAGAA 420 421 TCTTGTTCCTATCAATCTGGGGATCAAAAGCAAGCAATCCCACTACCGTGAA 480 421 TCTTGTTCCTATCAATCTGGGGATCAAAAGCAAGCAATCCCACTACCGTGAA 480 481 CCTAGGTTCTACACTAAGGCTCTCAAGAGGTGCTATTTCCAGATGATCTGGAACCTCT 540 481 CCTAGGTTCTACACTAAGGCTCTCAAGAGGTGCTATTTCCAGATGATCTGGAACCTCT 540 541 TATGAGCTCTCCCTACTAGCTGCAAAATGATTTGGAGCTTATGTCAGAAACCAAGAAAC 600 541 TATGAGCTCTCCCTACTAGCTGCAAAATGATTTGGAGCTTATGTCAGAAACCAAGAAAC 600 601 CAGAACCAAGATGATTTGGTTGGTGGCAAAAAGTTGATTTCAAGATGATGAAGGGTTCAG 660 601 CAGAACCAAGATGATTTGGTTGGTGGCAAAAAGTTGATTTCAAGATGATGAAGGGTTCAG 660 661 CGATTTCTGGATCTAGCCAGTACAGTGTAAATAGCAATCTCTGCGATATGAGGGGTATTT 720 661 CGATTTCTGGATCTAGCCAGTACAGTGTAAATAGCAATCTCTGCGATATGAGGGGTATTT 720 721 GGCCTCTGGTTATGTTAGCACTGAGGATATCAAAACCAAGAGGTTTGTGTCAATGCTG 780 721 GGCCTCTGGTTATGTTAGCACTGAGGATATCAAAACCAAGAGGTTTGTGTCAATGCTG 780									

QY	781	GACCTTGAGCCTGGCCAGACGGTCTCTGGATGTTGGTTGTGGAATTTGGTGGAGGTGACTTT	840
DB	781	GACCTTGAGCCTGGCCAGACGGTCTCTGGATGTTGGTTGTGGAATTTGGTGGAGGTGACTTT	840
QY	841	TACATGGCGGAGACCTTTGATGTTGAGGTTGTTGGATTTGATCTCTCCGTTAATATGATT	900
DB	841	TACATGGCGGAGACCTTTGATGTTGAGGTTGTTGGATTTGATCTCTCCGTTAATATGATT	900
QY	901	TCTTTGGCCCTTGAGCGTCTTATTTGGCTTAAATGCTGCTGCTGAGTTTGGAGTTCAGAT	960
DB	901	TCTTTGGCCCTTGAGCGTCTTATTTGGCTTAAATGCTGCTGCTGAGTTTGGAGTTCAGAT	960
QY	961	TGCACCAAGATAAATACCTACCTGATTAATCTTTTGAATGTCATCTATAGCCGTGACACATT	1020
DB	961	TGCACCAAGATAAATACCTACCTGATTAATCTTTTGAATGTCATCTATAGCCGTGACACATT	1020
QY	1021	CTGCATATTCAGGACCAAGCCTGCTGTTTGTAGATCTCTTCTACAAATGGTTGAAGCCAGCA	1080
DB	1021	CTGCATATTCAGGACCAAGCCTGCTGTTTGTAGATCTCTTCTACAAATGGTTGAAGCCAGCA	1080
QY	1081	GGTAAAGTTCTTAATCAGTGATTTAGTCTCAAGAAAGCTGGTCCACCTCACCTGAATTCGCC	1140
DB	1081	GGTAAAGTTCTTAATCAGTGATTTAGTCTCAAGAAAGCTGGTCCACCTCACCTGAATTCGCC	1140
QY	1141	GCTTACATTAAGCAGAGGGGATATGATCTTCCATGATGTAAAGGAATAATGGGCAGATGCTT	1200
DB	1141	GCTTACATTAAGCAGAGGGGATATGATCTTCCATGATGTAAAGGAATAATGGGCAGATGCTT	1200
QY	1201	AAAAGATGCTGCAATTTGTTGATGTTCTTGGCGAGGATAGAACTGAGCAGTTTCATTCGAGTT	1260
DB	1201	AAAAGATGCTGCAATTTGTTGATGTTCTTGGCGAGGATAGAACTGAGCAGTTTCATTCGAGTT	1260
QY	1261	CTACGGAAGAACTAGAGACTGTTGAGAGAAAGAGATGTTGTTCAATTTAGTATTTCTCT	1320
DB	1261	CTACGGAAGAACTAGAGACTGTTGAGAGAAAGAGATGTTGTTCAATTTAGTATTTCTCT	1320
QY	1321	GAGGAGGATTAACAATGCAATTTGTTGAGGTTGGAAATGATAAGTTGCGGAGGACTGCCAAG	1380
DB	1321	GAGGAGGATTAACAATGCAATTTGTTGAGGTTGGAAATGATAAGTTGCGGAGGACTGCCAAG	1380
QY	1381	GCTGAGCAACCTACCTGCTGTTGCTGCGAGGATAGAACTGAGCAGTTTCATTCGAGTT	1440
DB	1381	GCTGAGCAACCTACCTGCTGTTGCTGCGAGGATAGAACTGAGCAGTTTCATTCGAGTT	1440
QY	1441	GCCACTGTCGATTTCTCTAGTATTAATCTTCAATGTTTTCATGTAATGTAATTTCTACATGT	1500
DB	1441	GCCACTGTCGATTTCTCTAGTATTAATCTTCAATGTTTTCATGTAATGTAATTTCTACATGT	1500
QY	1501	AAAAATGGCAATAAGTTGCAATTTGCGAGACTGTAAGATGATTAATCAATTTTATCTTTT	1560
DB	1501	AAAAATGGCAATAAGTTGCAATTTGCGAGACTGTAAGATGATTAATCAATTTTATCTTTT	1560
QY	1561	AATTAATCATCGATTTATGCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA	1602
DB	1561	AATTAATCATCGATTTATGCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA	1602

## RESULT 3

AB080186	AB080186	1779 bp	mRNA	linear	PLN 20-FEB-2003
LOCUS					
DEFINITION	Suaeda japonica PEAMT mRNA for phosphoethanolamine N-methyltransferase, complete cds.				
ACCESSION	AB080186				
VERSION	AB080186.1	GI:28436073			
KEYWORDS					
SOURCE	Suaeda japonica				
ORGANISM	Suaeda japonica				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales; Amaranthaceae; Suaeda.				
REFERENCE	1				
AUTHORS	Yamada, A., Nozawa, G.T., Tanimoto, S. and Ozeki, Y.				
TITLE	Glycinebetaine synthesis in Suaeda japonica				

JOURNAL REFERENCE AUTHORS TITLE JOURNAL	Unpublished 2 (bases 1 to 1779) Yamada, A., Nozawa, G.T. and Ozeki, Y. Direct Submission Submitted (19-FEB-2002) Akiyo Yamada, Tokyo University of Agriculture and Technology, Department of Biotechnology; Naka-cho 2-24-16, Koganei, Tokyo 184-8588, Japan (E-mail: yamadenec.tuat.ac.jp. Tel.81-423-88-7239, Fax:81-423-88-7239)	
FEATURES	Location/Qualifiers 1. .1779 /organism="Suaeda japonica" /mol_type="mRNA" /db_xref="taxon:90346" /country="Japan: Soga, Ariake Sea" 1. .1779 /gene="PEAMT" 115. .1599 /gene="PEAMT" /codon_start=1 /product="phosphoethanolamine N-methyltransferase" /protein_id="BAC57432.1" /db_xref="GI:28436074" /translation="MAASGMELQBERDVPKQVWVHVDLTIEAMWLDQSADLDKEE RPEILSMPLPEKCHLELGLGIGRTGELAERKAGOVIALDFIESAIKKNVINGHYK NVKFMCAVTSFTUSFEPSHSDVIFSNWLLWYLSDEVENLVERMLKXPGVIFER ESCFQSGDHKKRNPHYREPRPTKAFKCHLQDGSNTELSLCKKICGAYEN KKNQHSIWLKQKVDSDKDGQFRLDTSQYKCNLSILRYRVPFGVYVETGYSITKE FVSLDLKPGQKVLVDCGIGGGDFYMAETFDVEVPDLSVNMISFALERSIIGLKA VEFEVADCTKINYPDNSFDVYSRDTILHIQDPALFRSFKMLKPGKVLISDYCKK AGPPSPFPAAYIKQGYDLHVKYQMLKADGFVDVLAEDRETOPFIRVLRKELETVE KEKDVFIQDFSEEDYNDIVGMDNKLRLRTAKGRWGLFVANKK"	
source		
gene		
CDS		
ORIGIN	Query Match 100.0%; Score 1602; DB 8; Length 1779; Best Local Similarity 100.0%; Pred. No. 0; Matches 1602; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Oy	1	CACACCGTGTGATTAACCAATGAAGCTATGATGCTCGAATCTCAAGCTTCTGATCTTGAC 60
Db	178	CACACCGTGTGATTAACCAATGAAGCTATGATGCTCGAATCTCAAGCTTCTGATCTTGAC 237
Oy	61	AAAGAGAAGCGTCTGAGATCTTTCAATGCTCCGCTCTTGAAGGAAAATGCTCTTG 120
Db	238	AAAGAGAAGCGTCTGAGATCTTTCAATGCTCCGCTCTTGAAGGAAAATGCTCTTG 297
Oy	121	GAATCTGGGCTGGATGCTGCTTTACTGCTGATGCTGCTGAGAGAGCTGCGCAGGTT 180
Db	298	GAATCTGGGCTGGATGCTGCTTTACTGCTGATGCTGCTGAGAGAGCTGCGCAGGTT 357
Oy	181	ATTGCTCTGATTTCAATGAGAGTCTATCAAGAGAAATGAAGTAATCAATGGGCACTAC 240
Db	358	ATTGCTCTGATTTCAATGAGAGTCTATCAAGAGAAATGAAGTAATCAATGGGCACTAC 417
Oy	241	AAAAATGTCAGTTTATGCTGCTCATGAGCTCTCCACCTCTCAGTTTCCACCAACAT 300
Db	418	AAAAATGTCAGTTTATGCTGCTCATGAGCTCTCCACCTCTCAGTTTCCACCAACAT 477
Oy	301	TCATFGAATGATATCTCCAAATGGTTACTCATGATCTTTCTGATGAAGAGGTGGAA 360
Db	478	TCATFGAATGATATCTCCAAATGGTTACTCATGATCTTTCTGATGAAGAGGTGGAA 537
Oy	361	AAATTTGGTTGAAAGATGTTGAATGGTTGAAGCCAGGGGCTTACATTTCTTCAGAGAA 420
Db	538	AAATTTGGTTGAAAGATGTTGAATGGTTGAAGCCAGGGGCTTACATTTCTTCAGAGAA 597
Oy	421	TCCTGTTTCCATCTCGGGATCACAAACCAAGCAATCCACCCACTACCGTGA 480
Db	598	TCCTGTTTCCATCTCGGGATCACAAACCAAGCAATCCACCCACTACCGTGA 657
Oy	481	CCTAGGTTCTACACATAAGGCTTTCAAGAGTGTCAATTTGCAAGATGATCTGGAACCTCT 540
Db	658	CCTAGGTTCTACACATAAGGCTTTCAAGAGTGTCAATTTGCAAGATGATCTGGAACCTCT 717

Oy	541	TATGAGCTCTCCCTACTAGCTGCAAAATGTATTTGGAGCTTATGTGAGAAACAAGAAAC 600
Db	718	TATGAGCTCTCCCTACTAGCTGCAAAATGTATTTGGAGCTTATGTGAGAAACAAGAAAC 777
Oy	601	CAGAACCAAGATAGTTGGTGTGGCAAAAGTTGATCTTAAGGATGATAAGGGGTTCAG 660
Db	778	CAGAACCAAGATAGTTGGTGTGGCAAAAGTTGATCTTAAGGATGATAAGGGGTTCAG 837
Oy	661	CGATTTCTGATACTAGCCAGTACAAGGTGTAATAGCAATCTCGGATATGAGCGTGTATT 720
Db	838	CGATTTCTGATACTAGCCAGTACAAGGTGTAATAGCAATCTCGGATATGAGCGTGTATT 897
Oy	721	GGCCCTGTTTATGTTAGCACTGAGAGATATGAACCAACCAAGAGTGTGTCATGCTG 780
Db	898	GGCCCTGTTTATGTTAGCACTGAGAGATATGAACCAACCAAGAGTGTGTCATGCTG 957
Oy	781	GACTTGAAGCCCTGGCCAGAGGTCTGGATGTTGGTGTGGTGTGGATTTGGTGGAGTCT 840
Db	958	GACTTGAAGCCCTGGCCAGAGGTCTGGATGTTGGTGTGGTGTGGATTTGGTGGAGTCT 1017
Oy	841	TACATGCGGAGACCTTTGATGTTGAGTGTGGATTTGATCTCTCCGTTTAATGATATT 900
Db	1018	TACATGCGGAGACCTTTGATGTTGAGTGTGGATTTGATCTCTCCGTTTAATGATATT 1077
Oy	901	TCCTTTGGCCCTTGAGCGTCTTATTTGGGCTTTAAATGTGCTTTGAGTTTGAGGTAGCAGAT 960
Db	1078	TCCTTTGGCCCTTGAGCGTCTTATTTGGGCTTTAAATGTGCTTTGAGTTTGAGGTAGCAGAT 1137
Oy	961	TGCAACCAAGATAAACTACCTGATTAACCTTTTGAATGTCATCTATAGCGCTGACACCAT 1020
Db	1138	TGCAACCAAGATAAACTACCTGATTAACCTTTTGAATGTCATCTATAGCGCTGACACCAT 1197
Oy	1021	CTGCATATTTCAGGCAAGCCCTGCGTCTGTTAGATCTCTTACAAATGGTTGAAGCCAGCA 1080
Db	1198	CTGCATATTTCAGGCAAGCCCTGCGTCTGTTAGATCTCTTACAAATGGTTGAAGCCAGCA 1257
Oy	1081	GCTTAAGTTCTAATCAGTATTAATGCAAGAGAGTGTGCTCCACCTCACTCAATTTGCGC 1140
Db	1258	GCTTAAGTTCTAATCAGTATTAATGCAAGAGAGTGTGCTCCACCTCACTCAATTTGCGC 1317
Oy	1141	GCTTACATTAAGCAGAGGGGATGATCTCCATGATGTAAGGAATATGGGCGAGTCTT 1200
Db	1318	GCTTACATTAAGCAGAGGGGATGATCTCCATGATGTAAGGAATATGGGCGAGTCTT 1377
Oy	1201	AAAGATGCTGGAATTTGATGTTCTTCCGAGGATAGAACTGAGCAGTTCATTTCCAGTT 1260
Db	1378	AAAGATGCTGGAATTTGATGTTCTTCCGAGGATAGAACTGAGCAGTTCATTTCCAGTT 1437
Oy	1261	CTACGAGGAACTAGAGACTGTTGAGAGGAAAGGATGCTGTTCAATGATGATTTCTCT 1320
Db	1438	CTACGAGGAACTAGAGACTGTTGAGAGGAAAGGATGCTGTTCAATGATGATTTCTCT 1497
Oy	1321	GAGGAGGATTAACATGACATTTGTTGGAGGTTGGAATCATTAAGTTGCGGAGGACTCGCAG 1380
Db	1498	GAGGAGGATTAACATGACATTTGTTGGAGGTTGGAATCATTAAGTTGCGGAGGACTCGCAG 1557
Oy	1381	GGTGACCAACGATGGGCTCTGTTCTGCAAGAGAGTGAAGCAATCAGTTCCGCACT 1440
Db	1558	GGTGACCAACGATGGGCTCTGTTCTGCAAGAGAGTGAAGCAATCAGTTCCGCACT 1617
Oy	1441	GGCACTGTGCGATTTCTAGTATTAATCTTCAATGTTTTCATGATGATCTTCTACATGT 1500
Db	1618	GGCACTGTGCGATTTCTAGTATTAATCTTCAATGTTTTCATGATGATCTTCTACATGT 1677
Oy	1501	AAAAATGCGCAATGAAGTTGCAATTTCCGACACTGTAAGATGATTAATCATATTTTATCTTT 1560
Db	1678	AAAAATGCGCAATGAAGTTGCAATTTCCGACACTGTAAGATGATTAATCATATTTTATCTTT 1737
Oy	1561	AAATTAATCATGGAATTTATGCAAAAAAATAAAAAAATAAAAAA 1602
Db	1738	AAATTAATCATGGAATTTATGCAAAAAAATAAAAAAATAAAAAA 1779





```

VERSION AF237633.1 GI:7407188
KEYWORDS Spinacia oleracea (spinach)
SOURCE Spinacia oleracea
ORGANISM Eukaryota; viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Amaranthaceae; Spinacia.
REFERENCE 1 (bases 1 to 2252)
AUTHORS Nuccio, M.L., Ziemak, M.J., Henry, S.A., Weretilnyk, E.A. and
Hanson, A.D.
TITLE cDNA cloning of phosphoethanolamine N-methyltransferase from
spinach by complementation in Schizosaccharomyces pombe and
characterization of the recombinant enzyme
J. Biol. Chem. 275 (19), 14095-14101 (2000)
MEDLINE 20361526
PUBMED 10799484
REFERENCE 2 (bases 1 to 2252)
AUTHORS Nuccio, M.L. and Hanson, A.D.
TITLE Direct Submission
JOURNAL Submitted (21-FEB-2000) Horticultural Sciences, University of
Florida, P.O. Box 110690, Gainesville, FL 32611, USA
FEATURES
source
1..2252
/organism="Spinacia oleracea"
/mol_type="mRNA"
/db_xref="taxon:3562"
/tissue_type="salinized leaf"
1..2252
/gene="PRMT"
254..1738
/gene="PRMT"
/EC_number="2.1.1.103"
/function="catalyzes S-adenosylmethionine dependent
N-methylation of phosphoethanolamine,
phosphonomethylmethanolamine and
phosphodimethylmethanolamine"
/codon_start=1
/product="phosphoethanolamine N-methyltransferase"
/protein_id="AA661950.1"
/db_xref="GI:7407189"
/translators="MAASANGVLOREVFVKYWEHSVDLTVEAMMLDSQASDLDKVE
RPEVSLMPEYGVLELQAGLGRFTGELAKASQVIALDFIESVINKKNSINGHYK
NPKWPCADVTSPSLNIPNSVDIIFSNMLMTLSDEEVLVERMLKPKPGYIPFR
ESCFHQSQDNKSNPHTYRPFYTKIPEKCHMODDSNYSVELSICGCTGAYVKS
KKNQOISLWQKVDSEDDKGPORFIDSSQYKFNISLRVYRFGYGVSTGGLTETKE
FVSKLDLKPQKVLQVCGGIGGGDFYMAENYDVEVVIDLSINMLKPKGKVLISDYCKS
VEFVACTKQKVPENSPVIYSRDTLHLOKPAKLPFSFHNKTDQPIQVLKELDALE
AGTFSABFAAYIRGQYDLHDVAYGKMLKDGAFYEVIAENRTDQPIQVLKELDALE
GEKDDFIDDFSEEDINDIVDGMKAKLVRTTEGGQWGLFIAKKM"

Query Match 71.1%; Score 1138.6; DB 8; Length 2252;
Best Local Similarity 82.6%; Pred. No. 4.9e-270;
Matches 1328; Conservative 0; Mismatches 274; Indels 5; Gaps 2;

QY 1 CACACGGTGTATTAACCATTAAGCTATGATGCTGATCTCAAGCTTCTGATCTTGAC 60
DB 317 CACTCTGTGTATGACTGTGTGAGGCTATGATGCTGTGATTCACAGCTTCAGATTTGAC 376
QY 61 AAAGAAGACGGCTCTGAGATCTTTTCAATGCTTCGGCTCTTCAAGAGAAATGCTCTTG 120
DB 377 AAATGGAGCGACCTGAGTACTTTCCATGCTTCCACCTTATGAGGAAAGTCTGTCTTA 436
QY 121 GAATCGGGCTGGTATTGCTGCTTTTACTGGTGAATTTGGCTGAGAAAGCTGGCCAGGTT 180
DB 437 GAATCGGGCTGGTATTGCTGCTTTTACTGGTGAATTTGGCCGAGAAAGCTAGCAGGTC 496
QY 181 ATTCCTCTGATTTCAATGAGAGTCTATCAAGAGAAATCAAGTATCAATGAGGACTAC 240
DB 497 ATCCCTCTGATTTCAATGAGAGTCTATCAAGAGAAATCAAGTATCAATGAGGACTAC 556
QY 241 AAAAATGTCAAGTTTATGTGCTGATGTGACTTCTCCCACTCTCAGTTTCCACCACAT 300

```

```

DB 557 AAAAATGTGAAGTTTATGTGTGCTGATGTGACATCTTCAAGTCTCAACATTTCCACCAAT 616
QY 301 TCATTGGATGTGATATTTCTCCAAATGGTTTACTCATGTATCTTTCTGTGATGAAGGTGGAA 360
DB 617 TCCGTGATATCATATTTCTCCAAATGGTACTCATGTATCTTTCTGTGATGAAGGTGGAG 676
QY 361 AATTGGTTCAGAAAGATTTGAAATGGTTGAAGCAGGGGGTTACATTTCTTTCAGAGAA 420
DB 677 CGTCTGGTTGAAGGATTTGAAATGGTTGAAGCAGGGAGATACATTTCTTTCAGAGAA 736
QY 421 TCTTGTTCATCATCATCTGGGGATCAAAACCAAGCAATCCACCACCTACCGTGA 480
DB 737 TCTTGTTCATCATCAGAGATCAAGCCGCAAAAGCAATCCAAACCACTACCGTGA 796
QY 481 CCTAGGTTCTACACTAAGGCCCTTCAAGAGTGTCAATTTGCAAGATGATGTTGGAATCTCT 540
DB 797 CTTAGGTTCTACACCAAGATCTTCAAGAAATGCCATATGCAAGATGATTTCTGGCACTCC 856
QY 541 TATGAGCTCTCCCTACTAGCTGCAATCTATTTGGAGCTTATGTCAGAAACAAGAAAC 600
DB 857 TATGAGCTCTCCCTTATTTGGCTCCAAATGTTATTTGGAGCTTATGTCAGAAACAAGAAAT 916
QY 601 CAGAACAGATTTAGTTGGTTGTCGCAAAAGTTGATTTCAAGATGATAAGGGGTTCAG 660
DB 917 CAGAACAGATTAAGCTGTTTATGCGCAGAAAGTTGATTTCAAGAGATGACAAAGGGGTTCAG 976
QY 661 CGATTTCTGGATCTAGCCAGTACAAGTGTAAATAGCAATCTCGCATATAGAGGTGATTT 720
DB 977 CGATTTCTGGATCTAGCTCAATACAGTTTACAGCTACTGCGTTATGAGGGGTGATTT 1036
QY 721 GGCCTCTGTTATGTTAGCACTGGAGGATATGAACCAACCAAGAGTTTGTGTCAATGCTG 780
DB 1037 GGTCTCTGTTATGTTAGTACCGAGGACTCGAAACCAACCAAGAGTTGTTGATCAAGAGTT 1096
QY 781 GACTTTGAAGCCTGCGCAGAAAGTCTCGGATGTTGGTGTGTTGGAAATTTGGTGGAGTCACTTT 840
DB 1097 GACTTTGAAGCCTGCGCAGAAAGTCTTAGATGTGGTGTGTTGGCATAGTGGAGGTGATTT 1156
QY 841 TACATGGCGAGACCTTTGATGTTGAGTGTGTTGGAATTTGATCTCTCCGTAAATATGAT 900
DB 1157 TACATGGCGAGAGAACTATGATGTTGAGTGTGTTGGAATTTGATCTCTCCGTAAATATGAT 1216
QY 901 TCTTTTGGCTTTGAGCGTTCTATTGGGCTTAAATGCTGTTGAGTGTGAGTGTGAGTAGCAGAT 960
DB 1217 TCTTTTGGCTTTGAGCGCTCAATTTGGCTCAAAATGCTGTTGAGTGTGAGTGTGAGTAGCAGAT 1276
QY 961 TGCACCAAGATAAACTACCTGATAACTCTTTTGTGATGTCATCTATAGCCGTGACACCAAT 1020
DB 1277 TGCACCAAGAAAGATTACCTTGAAACTCTTTTGTGATGTCATCTATAGCCGTGATACCAAT 1336
QY 1021 CTGCAATTTCAAGACAGCCCTGCTGTTGTTAGATCTCTTCAAAATGTTGAGCCAGCA 1080
DB 1337 CTGCAATTTCAAGACAAACCTGCTTTTATTTAGATCTCTTCAAAATGTTGAAACCTGGA 1396
QY 1081 GCTAAAGTTCTTAATCAGTGAATTACTGCAAGAAAGCTGGTCCACCTCACTCAATTCGCC 1140
DB 1397 GCTAAAGTTCTTATTTAGTACTCTGTAAGAGTCTGGTACACTTCAGCTGAATTTGCT 1456
QY 1141 GCTTACATTTAGCAGAGGGGATATGATCTCCATGATGTAAAGCAATATGGCAGAGTCTT 1200
DB 1457 GCATACATCAGCAGAGGGGATATGATCTCCATGATGTAAAGCAATATGGCAGAGTCTT 1516
QY 1201 AAAGATGTGGATTTGATGTTGTTGCTTGGCCAGAGATAGAACTCAGCAGTTCATTCAGATT 1260
DB 1517 AAAGATGTGGATTTGATGTTGTTGCTTGGCCAGAGATAGAACTCAGCAGTTCATTCAGATT 1576
QY 1261 CTACCGAGAGGACTAGAGACTGTTGAGAGGAAAGGATGTTTCAATTTAGTGTGATTTCTCT 1320
DB 1577 CTGACAGAGGACTAGAGTCTTGAACAGAGAGGATGATCTTCAATTTAGTGTGATTTCTCT 1636
QY 1321 GAGAGGATTAACATGACTTTTGGAGGTTTGGAGTGTGAATGATTAAGTTCGGAGGACTGCAAG 1380
DB 1637 GAGAGGATTAACATGACTTTTGGAGGTTTGGAGTGTGAATGATTAAGTTCGGAGGACTGCAAG 1696

```



QY	1201	AAAGATGCTGGATTGTTGATGTTCTTCCGAGGATAGAACTGAGCAGTTCATTCCAGTT	1260
DB	1360	AGAGATGCTGGATTGTTGATGTTCTTCCGAGGATAGAACTGAGCAGTTCATTCCAGTT	1419
QY	1261	CTACGGAAGAACTAGAGACTGTTGAGAGAAAGGATGTTTCATTAGTATTTCTCT	1320
DB	1420	CTGCAAAAGGATTTAGATATCTGCGAGAGAAAGGATTTTATATCCATGAGTTCTCT	1479
QY	1321	GAGGAGGATTAACAATGACATGTTGAGGCTTGGATGATGAATGTTGCGAGAGACTGCCAAG	1380
DB	1480	GAACAGACTATATGATGATGTTGAGGCTTGGATGATGAATGTTGCGAGAGACTGCCAAG	1539
QY	1391	GCTGAGCAAGATGGGCTGTTGCTGTTGCGAAGAGAGAGTGAAGTCAAGTTCGCGACT	1440
DB	1540	GCTGAGCAAGATGGGCTGTTGCTGTTGCGAAGAGAGAGTGAAGTCAAGTTCGCGACT	1599
QY	1441	GGCACTGTCGATTTCTCTAGTATTAATCTTCAAGTGTTCATGTA	1495
DB	1600	GGATCTATTTCCATTTATATATCTAGAGTATGAACATTTCTTCAGTCTTATGTAATAT	1659
QY	1496	CATGTAAATTTGCCATTAAGTTCGATTTGCGAGACTGTAAAGTATGAATTAATATTTAT	1555
DB	1660	GTTTACCAAGGATTTAGGATTAAGATTAATTAAGTATTAAGTATTAAGTATTAAGT	1719
QY	1556	CTTTTAAATTAATCATGATTTATGCAAAAAAAGAAAAAAGAAAAAAGAAAAA	1602
DB	1720	AAATTTGTTCTTCTTCTGCTGTCAAAAAAGAAAAAAGAAAAAAGAAAAA	1766
RESULT 7			
AB090883			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
CDS			

EFVADACTKSYDPSFVIVSRDILHIQDKPALFRTFKMLKPGCKVLISDYCKKS			
GKPSDFABYIKORGYDLHVBVTYKMLKXGAGFVIAEDRTKQFVLEVLRELRVERK			
EKKEPTQDPTGDIYNDIVGGWAKLVRTSGEORWGLFIAROKK			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 1078; Conservative			
53.5%; Score 857.4; DB 8; Length 1870;			
75.5%; Pred. No. 1.2e-200;			
0; Mismatches 346; Indels 3; Gaps 1;			
QY	1	CACACGCTGATTTAAACCACTTGAAGCTATGATGCTGCAATCTCAAGCTTCTGATCTTGAC	60
DB	248	CACCTGCTGATCTTAACTGTTGAATCTATGATGCTTGAATCAATGCTTCTGATCTTGAT	307
QY	61	AAAGAGAAAGCTCTCTGAGATTTCTTCAATGCTTCCGCTCTTGAGAGAAATGCTCTTG	120
DB	308	AAAGAGAAAGCTCTCTGAGATTTCTTCAATGCTTCCGCTCTTGAGAGAAATGCTCTTG	367
QY	121	GAACCTGGGCTGATTTGCTGCTTTTACTGCTGAATTTGGCTGAGAAAGCTGGCCAGATT	180
DB	368	GAACCTGGGCTGATTTGCTGCTTTTACTGCTGAATTTGGCTGAGAAAGCTGGCCAGATT	427
QY	181	ATTCCTCTGATTTCACTGAGAGTCTATCAAGAGAAATGAGTAACTCAATGGGCATAC	240
DB	428	ATAGCCCTTGATTTCACTGAGAGTCTATCAAGAGAAATGAGTAACTCAATGGGCATAC	487
QY	241	AAAAATGTCAGATTTATGCTGCTGATGCTTCTCCACTCTCAGTTTCCCAACAT	300
DB	488	AAAAATGTCAGATTTATGCTGCTGATGCTTCTCCACTCTCAGTTTCCCAACAT	547
QY	301	TCATGAGATGATTTCTCCAACTTCTTACTCATGATCTTCTTCTGATGAGAGGTGAA	360
DB	548	TCATGAGATGATTTCTCCAACTTCTTACTCATGATCTTCTTCTGATGAGAGGTGAA	607
QY	361	AAATTTGGTTGAAAGATTTGAAATGTTGAGCCAGGGGGTTTACATTTCTTCAGAGAA	420
DB	608	GATATTCGGGAAAGGTTTTTGAATGCTGAGGTGGAGGACATATTTCTTCAGAGAG	667
QY	421	TCCTGCTTCCCAATCAATCTGGGATCAAAAGCGAAAGCAATCCCAACCACTACCTGAA	480
DB	668	TCATGCTTTCCCAATCTGGGATCAAAAGCGAAAGCAATCCCAACCACTACCTGAA	727
QY	481	CTAGGCTTCTACATAGGCTTCAAGAGTGTCTTTCGAGATGATGATCTGGAATCT	540
DB	728	CCTAGATTTCTACCAAGGCTTTCAGCAATGTCATGACCATGATCTGGAATCT	787
QY	541	TATGAGCTCTCCCTACTAGCTGCAAAATGTAATGGAGCTTATCTCAGAAACCAAGAAAC	600
DB	788	TATGAGCTCTCCCTACTAGCTGCAAAATGTAATGGAGCTTATCTCAGAAACCAAGAAAC	847
QY	601	CAGAACCAAGTATGTTGCTGCTGCAAAAGTTG---ATTCTAAGATGATTAAGGGCTTC	657
DB	848	CAGAACCAAGTATGTTGCTGCTGCAAAAGTTG---ATTCTAAGATGATTAAGGGCTTC	907
QY	658	CAGCGATTTCTGATATCTAGCAGTACAGTGTATAGCAATTTCTGATATGAGCTGTA	717
DB	908	CAGCAATTTCTGATTAATGGCCAGTACAAAGCAGTGGCATTTCTACGCTATGAGCGATA	967
QY	718	TTTGGCCCTGCTGTTATGTTAGCACTGGAGGATATGAAACCAACCAAGAGTGTGTCT	777
DB	968	TTTGGCCCTGCTGTTATGTTAGCACTGGAGGATATGAAACCAACCAAGAGTGTGTCT	1027
QY	778	CTGAGCTTGAAGCTCTGCGCAGAGAGTCTGATGCTTGGTGTGGAATGCTGAGGTGAC	837
DB	1028	CTTGAATTTAAAGCCCGCCAGAAAGTTTTGGATGTTGATGTTGATTTGGCATTTGAT	1087
QY	838	TTTATCATGCGCGAGACCTTTTGAATGTTGAGGTGTTGATTTGATTTCTCTCGCTTAT	897
DB	1088	TTCTATATGCGAGATAAATCTCGATGTTGATTTGATTTGATTTCTCTCTGTAATATG	1147
QY	898	ATTTCCTTTGCTTCTGAGGCTTCTTATGCGCTTAAATGCTGTTGATTTGAGTTGAGCA	957
DB	1148	ATAGCTTTTCTCTTCTGAAAGATCAATTTGGCTTCAATTTCTCAGTGGAAATTCAGGTT	1207

958 GATTCCACCAAGATAAATACCTACCTGATPACTCTTTTGTGTCTATCTATAGCGGTGACACC 1017  
 Db GACTGTACTAAGAAATCATATCCGGACAATTCGTTTGTGTCTATCTATAGCGGTGACACC 1267

1018 ATTCTGCATATTCAGGACCAAGCTGCGTTGTTTGTAGATCTCTACAAATGTTGAAGCCA 1077  
 Db ATCTCTCCATATCCAGGATAAAGCTGCTTTGTTCCGAACCTTCTACAGTGTGTTGAAGCCA 1327

1078 GGAGGTAAAGTTCTAAATCACTGATATCTGCAAGAAAGCTGCTCCACCTCACTGTAATTC 1137  
 Db GTGGCAAGTTCTCATTTAGTATCTGCGGAAAATCTGGAAAACCGTCTGAAGATTTT 1387

1138 GCCGTTTACATTAAGCAGAGGGATATGATCTCCATGATGTAAAGGAATATGGGCAAGATG 1197  
 Db GCAGGTATCATCAAGCAAGAGGATATGATCTCCATGATGTTTGAAGAAATATGGCAAGATG 1447

1198 CTTAAAGATGCTGCAATTTGTTGATGTTCTTGGCGAGGATAGACATGACGATCTCAATCGA 1257  
 Db CTTAAAGATGCTGCGCTTTGGAGAGGTGATAGCAGAGGATCTGTAACAAACGTTTCAATGAA 1507

1258 GTTCTACGAGGAACTAGACATGTTGAGAGGAAAGAGATGTTTCAATGATGATTC 1317  
 Db GTCTCTACGAGGAACTGGAAGAGTTGAGAAAGAAAGAGATTTCAATCCAGACTTC 1567

1318 TCTCAGGAGGATACATATGATGATGTTTGGAGGTTGCAATGATGATGCGGAGGACTGCC 1377  
 Db ACTGAAGTGACTACACAGATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1627

1378 AAGGTGAGCAAGATGCGGTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1424  
 Db TCTGGGAGCAGATGCGGTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1674

RESULT 8  
 AF428454  
 LOCUS  
 DEFINITION  
 Arabidopsis thaliana At1g48600/T1N15\_20 mRNA, complete cds.  
 ACCESSION  
 AF428454  
 VERSION  
 AF428454.1 GI:16226648  
 KEYWORDS  
 FLI CDNA.  
 SOURCE  
 Arabidopsis thaliana (thale cress)

ORGANISM  
 Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopses.  
 1 (bases 1 to 1656)  
 Chou, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Ban, J.,  
 Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,  
 Iehida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G.,  
 Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M.,  
 Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K.,  
 Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C.,  
 Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K.,  
 Davis, R.W., Theologis, A. and Ecker, J.R.  
 Arabidopsis  
 Unpublished  
 2 (bases 1 to 1656)  
 Chou, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Ban, J.,  
 Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,  
 Iehida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G.,  
 Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M.,  
 Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K.,  
 Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C.,  
 Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K.,  
 Davis, R.W., Theologis, A. and Ecker, J.R.  
 Arabidopsis  
 Unpublished

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 1 (bases 1 to 1656)  
 Chou, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Ban, J.,  
 Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,  
 Iehida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G.,  
 Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M.,  
 Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K.,  
 Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C.,  
 Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K.,  
 Davis, R.W., Theologis, A. and Ecker, J.R.  
 Direct Submission  
 Submitted (05-OCT-2001) Salk Institute Genomic Analysis Laboratory  
 (SIGAL), Plant Biology Laboratory, The Salk Institute for  
 Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,  
 USA.

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 2 (bases 1 to 1656)  
 Chou, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Ban, J.,  
 Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,  
 Iehida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G.,  
 Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M.,  
 Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K.,  
 Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C.,  
 Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K.,  
 Davis, R.W., Theologis, A. and Ecker, J.R.  
 Direct Submission  
 Submitted (05-OCT-2001) Salk Institute Genomic Analysis Laboratory  
 (SIGAL), Plant Biology Laboratory, The Salk Institute for  
 Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,  
 USA.

Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,  
 Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PISC (SSP) Consortium members carried out the  
 sequencing and annotation of the RAFL cDNAs: Cheuk, R., Chen, H.,  
 Kim, C.J., Koesema, E., Meyers, M.C., Shinn, P., Ban, J., Bowser, L.,  
 Dale, J.M., Goldsmith, A.D., Jiang, P.X., Jones, T., Karlin-Neumann, G.,  
 Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M.,  
 Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Southwick, A.,  
 Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S.,  
 Davis, R.W., Theologis, A., and Ecker, J.R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to  
 this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)  
 contributed equally to this work as PIs.

FEATURES  
 source

1. .1656  
 /location/Qualifiers  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /chromosome="1"  
 /clone="RAPL09-71-K18 (R14923)"  
 /ecotype="Columbia"  
 1. .18  
 19. .1446

5'UTR  
 CDS

/notes="putative phosphoethanolamine N-methyltransferase"  
 /codon\_start=1  
 /product="At1g48600/T1N15\_20"  
 /protein\_id="AAL16223.1"  
 /db\_xref="GI:16226649"  
 /translation="MEHSSDLTVAMMLDSKASDLKDEPEREVLISLPYEGKSVLEL  
 GAGIRFTCLAQAGVETADLPESATOKNESVNGHYKTFKPCADVTSPDLKIDG  
 SIDLIPSNLLAVLSDKEVELMAEIKGWAPFGTTPFRCFSGDSGRKSNPHTY  
 REPFYTKVFOECQIRDSAGSLFSGVCKGCIKAYVKNKNQNCIWIWQKVSVD  
 KDFQRLDNDVQYKSSGLIRYFVGVSTGGFTTKEFVKAMDLPKQKVLVDVGG  
 IGGDSYMAENFDVHVGLSDLSNMISFALERAIGLKSVEFVADCTTKTTPNSD  
 VYISRDITLHQDKPALFTFVKLPGKQKVLITDYCSAETSPSPAETIKQKGYDL  
 HDVQYQMLKADGDFDVIADDTQFQVLRLEKEVEKEEFISDFSEEDYNDIV  
 GWSAKLERTASGEQKMGFLIADKK"  
 1447. .1656

3'UTR  
 ORIGIN

Query Match 53.1%; Score 853.4; DB 8; Length 1656;  
 Best Local Similarity 75.2%; Pred. No. 1.2e-199;  
 Matches 1064; Conservative 0; Mismatches 351; Indels 0; Gaps 0;

QY 9 TGATTTAAACATTTGAAGCTATGATGCTCGATTTCTCAGCTTCTGATCTTGACAAAGAAGA 68  
 Db 33 TGATTTGACTGTTGAAGCTATGATGCTTGACTCTAAAGCTTCTGATCTTGACAAAGAAGA 92

QY 69 ACGTCTCTGAGATCTTTCAATGCTTCCGCTCTTGAAGGAAATGCTCTTGGAAGCTGG 128  
 Db 93 ACGTCTCTGAGTACTCTCTTTAATCCACCATATGAGGAAATCTGCTTGACTGG 152

QY 129 GCGTGTATTTGCTGCTTTTACTGTTGTAATGGCTGAGAAAGCTGCCAGGTTATGCTCT 188  
 Db 153 AGCTGGTATTTGGTCTGTTTCACTGGTGAATTTGGCTCAAAAGGCTGGTGAAGTTATGCTCT 212

QY 189 GGATTTCAATTCAGAGTGTCTATCAAGAAAGATGAGTAAATCAATGGGCACTACAAATATG 248  
 Db 213 TGATTTCAATTCAGAGTGTCTATCAAGAAAGATGAGTAAATCAATGGGCACTACAAATATG 272

QY 249 CAAGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 308  
 Db 273 CAAGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 332

QY 309 TGTGATATTTCTCAATTTGGTACTCATGTATCTTTCTGATGAAGAGGTGAAAATTTGGT 368  
 Db 333 CTGATTTTCTCAACTGGTGTCTCATGTATCTCTCTGATGAAGAGGTGAACTATATGCT 392

QY 369 TGAAGAAATTTGAATGTTGTAAGCAGCGGGGTACATTTTCTTCTCAGAGATCTGTTT 428  
 Db 393 AGAGAGATTTGATTTGGATGGTTCAGCCAGGGGGATACATTTTCTTCTCAGAGATCTGTTT 452



429 CCATCAATCTGGGATCACAAACGCAAAAGCAATCCACCACTACCGTGAACCTAGCTT 488  
 453 CCATCAATCTGGGACAGCAGCAAGCAAGCAAAACCCCACTACCGTGAACCCAGCTT 512  
 489 CTACATCAAGCTCTCAAGAGTCTCTTTCGAGAGTCTGAGTCTGAGTCTTATGAGCT 548  
 513 CTACAAAGGTTTTCAGAGTCTCTGAGAGTCTGAGTCTGAGTCTTATGAGCT 572  
 549 CTCCTACTAGCTGCAATGTTATGAGTCTTATGAGTCTTATGAGTCTTATGAGCT 608  
 573 CTCTATGCTGCTGCAATGCTATGAGTCTTATGAGTCTTATGAGTCTTATGAGCT 632  
 609 GATTAGTCTGCTGCAATGCTATGAGTCTTATGAGTCTTATGAGTCTTATGAGCT 668  
 633 GATTAGTCTGCTGCAATGCTATGAGTCTTATGAGTCTTATGAGTCTTATGAGCT 692  
 669 GGATCACTAGCTGCAATGCTATGAGTCTTATGAGTCTTATGAGTCTTATGAGCT 728  
 693 GGATCACTAGCTGCAATGCTATGAGTCTTATGAGTCTTATGAGTCTTATGAGCT 752  
 729 TTATGTTAGCTGCAATGCTATGAGTCTTATGAGTCTTATGAGTCTTATGAGCT 788  
 753 ATATGTTAGCTGCAATGCTATGAGTCTTATGAGTCTTATGAGTCTTATGAGCT 812  
 789 GCTGTTAGCTGCAATGCTATGAGTCTTATGAGTCTTATGAGTCTTATGAGCT 848  
 813 ACCGTTAGCTGCAATGCTATGAGTCTTATGAGTCTTATGAGTCTTATGAGCT 872  
 849 GGATCACTAGCTGCAATGCTATGAGTCTTATGAGTCTTATGAGTCTTATGAGCT 908  
 873 TGATCACTAGCTGCAATGCTATGAGTCTTATGAGTCTTATGAGTCTTATGAGCT 932  
 909 CTTGAGCTGCTGCAATGCTATGAGTCTTATGAGTCTTATGAGTCTTATGAGCT 968  
 933 ACTGAGCTGCTGCAATGCTATGAGTCTTATGAGTCTTATGAGTCTTATGAGCT 992  
 969 GATCACTAGCTGCAATGCTATGAGTCTTATGAGTCTTATGAGTCTTATGAGCT 1028  
 993 CAATCACTAGCTGCAATGCTATGAGTCTTATGAGTCTTATGAGTCTTATGAGCT 1052  
 1029 TCGAGCTGCTGCAATGCTATGAGTCTTATGAGTCTTATGAGTCTTATGAGCT 1088  
 1053 CCAAGCTGCTGCAATGCTATGAGTCTTATGAGTCTTATGAGTCTTATGAGCT 1112  
 1089 TCACTAGCTGCTGCAATGCTATGAGTCTTATGAGTCTTATGAGTCTTATGAGCT 1148  
 1113 TCACTAGCTGCTGCAATGCTATGAGTCTTATGAGTCTTATGAGTCTTATGAGCT 1172  
 1149 TCACTAGCTGCTGCAATGCTATGAGTCTTATGAGTCTTATGAGTCTTATGAGCT 1208  
 1173 TCACTAGCTGCTGCAATGCTATGAGTCTTATGAGTCTTATGAGTCTTATGAGCT 1232  
 1209 TCACTAGCTGCTGCAATGCTATGAGTCTTATGAGTCTTATGAGTCTTATGAGCT 1268  
 1233 TCACTAGCTGCTGCAATGCTATGAGTCTTATGAGTCTTATGAGTCTTATGAGCT 1292  
 1269 TCACTAGCTGCTGCAATGCTATGAGTCTTATGAGTCTTATGAGTCTTATGAGCT 1328  
 1293 TCACTAGCTGCTGCAATGCTATGAGTCTTATGAGTCTTATGAGTCTTATGAGCT 1352  
 1329 TCACTAGCTGCTGCAATGCTATGAGTCTTATGAGTCTTATGAGTCTTATGAGCT 1388  
 1353 TCACTAGCTGCTGCAATGCTATGAGTCTTATGAGTCTTATGAGTCTTATGAGCT 1412  
 1389 TCACTAGCTGCTGCAATGCTATGAGTCTTATGAGTCTTATGAGTCTTATGAGCT 1423  
 1413 TCACTAGCTGCTGCAATGCTATGAGTCTTATGAGTCTTATGAGTCTTATGAGCT 1447

RESULT 9  
 AV133811  
 LOCUS

DEFINITION Arabidopsis thaliana clone U1999 putative phosphoethanolamine N-methyltransferase (At1g48600) mRNA, complete cds.

ACCESSION AV133811  
 VERSION 1  
 KEYWORDS FLI CDNA  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 REFERENCE 1 (bases 1 to 1459)  
 AUTHORS Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hsuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yuan, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.  
 Arabidopsis Open Reading Frame (ORF) Clones  
 Unpublished  
 2 (bases 1 to 1459)  
 TITLE Direct Submission  
 JOURNAL Submitted (17-JUL-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA  
 REFERENCE The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.  
 COMMENT The Salk, Stanford, PGEN (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAPL cDNAs: Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hsuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yuan, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.  
 YAMADA, K. (SSP/PGEN) and SEKI, M. (RIKEN GSC) contributed equally to this work. SHINOZAKI, K. (RIKEN GSC) and THEOLOGIS, A. (SSP/PGEN) contributed equally to this work as PIs.  
 Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.  
 FEATURES  
 source  
 1..1459  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /chromosome="1"  
 /clone="U11995"  
 /ecotype="Columbia"  
 /note="this clone is in PUNI 51."  
 1..1459  
 /gene="At1g48600"  
 1..1428  
 /gene="At1g48600"  
 /codon\_start=1  
 /evidence="experimental"  
 /product="putative phosphoethanolamine N-methyltransferase"  
 /protein\_id="AA091745.1"  
 /db\_xref="GI:22136802"  
 /translation="MEHSSDUTVEAMMLDSKASDLKSRPEVLSLIPPEGKSVLEL"





Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinohara, K., Davis, R.W., Ecker, J.R. and Theologis, A.  
 Direct Submission  
 Submitted (19-NOV-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA  
 RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinohara, K.  
 The Salk, Stanford, PGECC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada, K., Ban, H., J., Banno, F., Chang, E., Dale, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Koesema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shim, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGECC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinohara, K. (RIKEN GSC) and Theologis, A. (SSP/PGECC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

#### FEATURES

Location/Qualifiers  
 1..1660  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /chromosome="1"  
 /clone="RAFL06-86-D12 (R11999)"  
 /ecotype="Columbia"  
 /note="This clone is in a modified pbluescript vector (Lambda Zap) as a XhoI/SatI insert."  
 1..1660  
 /gene="At1g48600"  
 1..34  
 /gene="At1g48600"  
 35..1462  
 /gene="At1g48600"  
 /codon\_start=1  
 /evidence=experimental  
 /product="putative phosphoethanolamine N-methyltransferase"  
 /protein\_id="AA13622.1"  
 /db\_xref="GI:17360784"  
 /translation="MEHSDLTVEAMWLDKSKADLDKRRPEVLSLIPPGKSVLEL  
 GAGTGRFTGELAQAGHVIADLFESAIOQNSVNGHYKIKPMCAADVSPDLKIKD  
 SIDLIFNMLYLSKXVBLMAERIMGWKPGGYFFPESCFHQSGDSKRSNPHY  
 REPFRYTKVFCOTRDSNGSFGSCGICGAYVKNKQNOICWIKQSVEND  
 KDFORFLDVOYKSSGILYRVFEGYVSTGFTFVFVAKMDLKEGKVLVDGCG  
 LGGDFPMANFDFVGVGIDLSVNMISPAELGKCSVEFVADCTTKYTPNSFD  
 VTYSDTLHIQDKPALFRFPKMLPGKGLVLDYCRSAETPSPEFARYIKQRYDL  
 HDVQAGMLKDAQSPDVIADETDQPVQVRLSELEKEKEEPEISDFSEYNDIV  
 GNSALERTASGQKGLFIADKK"  
 misc\_difference 316  
 /gene="At1g48600"  
 /note="compared to genomic sequence"  
 /replace="t"  
 1463..1660  
 /gene="At1g48600"

#### ORIGIN

Query Match 53.2%; Score 851.8; DB 8; Length 1660;  
 Best Local Similarity 75.1%; Pred. No. 3e-199;  
 Matches 1063; Conservative 0; Mismatches 352; Indels 0; Gaps 0;  
 9 TGATTTACCATGAGCTATGATCTCGATCTCAAGCTTCGATCTTGACAAAGACA 68  
 49 TGATTTGACTGTGAGCTATGATCTTGACTCTTAAAGCTTCTGATCTGACAAAGACA 108

69 ACCTCTGAGATCTTTCAATGCTTCGCTCTTGAGGAAAAATGCTCTTGGAAGCTTGG 128  
 109 ACCTCTGAGGTAATCTCTTTAATCCACCATATAGGGAAATCTGTGCTTGGAAGCTTGG 168  
 129 GGCTGGTATTGGTCTGTTTACTGGTGAATGGCTGAGAAAGCTGGCCAGGTTATTGCTCT 188  
 169 AGCTGGTATTGGTCTGTTTACTGGTGAATGGCTCAAAAGGCTGGTGAAGTTATCGCTCT 228  
 189 GGATTTCAATGAGAGGCTATCAAGAGAGATCAAGTATCAATGGGCACTACAAAGATGT 248  
 229 TGATTTCAATGAAGCGGCAATCAGAGAGATCAAGTATGAGTATGGGCAATCAAGAGAT 288  
 249 CAAGTTATGCTGCTGATGATCTTCCACTCTCAGTTTCCACACACATCAATTCATGGA 308  
 289 CAAGTTATGCTGCTGATGATCTTCCACTCTCAGTTTCCACACACATCAATTCATGGA 348  
 309 TGATGATTTCCAAATGGGTTACTCATGATCTTCTCATGAAGAGGTGGAAAAATTTGGT 368  
 349 CTTCATTTCTCAAACTGGTTGCTCATGATCTCTCATGAAGAGGTGGAAAAATTTGGT 408  
 369 TGAAGATGTTGAATGTTGAGCCAGGCTTACATTTCTTCTCAGAGATCTTCTTCTT 428  
 409 AGAGAGATGATGGATGGGTCAAGCCAGGGGATACATTTCTTCTCAGAGATCTTCTTCT 468  
 429 CCATCAATCTGGGATCAAAAGCGAAAAAGCAATCCCACTACCTGAACTTAGGT 488  
 469 CCATCAATCTGGGATCAAAAGCGAAAAAGCAATCCCACTACCTGAACTTAGGT 528  
 489 CTCACTAAGGCTTCAAGAGGTCTCATTTGCAAGATGGATCTGAAATCTTTATGAGCT 548  
 529 CTACAAAGGTTTTCAGGATGTGACACATGCTTCTTCTGCAATTCATTTGAGCT 588  
 549 CTCCCTACTAGCTGCAATGATTTGAGCTTATGTGAGAAACAAGAAAAACCAAGACA 608  
 589 CTCTATGTTGGTGCATATGCAATGGGCTTATGTGAAGAAACAAGAAAAACCAAGACA 648  
 609 GATAGTTGGTGTGGCAAAAAGTTGATCTTAAGAGATGATAAGGGGTTCCAGCGATTTCT 668  
 649 GATTTGCTGATATGGCAAAAAGTTGAGGTTGAGAGATGACAAAGATTTCCAGCGTTCTT 708  
 669 GATTTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAG 728  
 709 GCAAAATGTTCAATCAAGTCTAGTGGGATCTTGGCTATGAGCGGTCTTTTGGGGAAGG 768  
 729 TTATCTTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAG 788  
 769 ATATGTGAGCACTGGTGGATTTGAGCAAACTAAAGAAATTTGTGGCAAGATGGACCTTAA 828  
 789 GCTGGCCAGAGCTCTGGATGTTGGTGGAAATTTGGTGGAGGTGACTTTTATCATGCG 848  
 829 ACCGGGACAGAGTCTTAGATGTTGGTGGTATCGGTGGAGGTGACTTTTATCATGCG 888  
 849 GGAGACCTTTGATGTTGAGGTTGATGATCTCTCGTTAATATGATTTCTTCTTGGC 908  
 889 TGAGAAATTCGATGTTCAATGTTTGGATCGATCTGTGGTCAACATGATCTCTTCTGCG 948  
 909 CCTTCAGGCTTCTATTTGGCTTAAATGCTGTGTGATGTTGAGTATGAGTATGACAGTTCAC 968  
 949 ACTGAGCGGGCATTTGAGCTCAATGCTCAGTGGATTTGAGTATGAGTATGAGTATGAG 1008  
 969 GATAAATCTGCTGATTAATCTCTTTTATGATGATCTATAGCGGTGACACCATTTCTGATAT 1028  
 1009 CAATAACATATCCGATAATCTCTTTTATGATGATCTATAGCGGTGACACCATTTCTGATAT 1068  
 1029 TCAGGACAGGCTGCTGTTGTTTATGATCTCTTCAATAATGTTGAAGCGGAGGTTAAAGT 1088  
 1069 CCAAGACAGCGAGCTCTATTTCAAGGATCTTCAAGTGGCTTAAACACGAGGGGTAAAGT 1128  
 1089 TCTAATCAGTGAATCTCAAGAAAGCTGGTCCACCTGCACTCTGAAATTCGCGCTTACAT 1148  
 1129 TCTAATCAGTGAATCTCAAGAAAGCTGGTCCACCTGCACTCTGAAATTCGCGCTTACAT 1188





QY 61 AAGAGAGAGCTCTGAGATCTTTCAATGCTTCCGCTCTTTGAGGAAATGCCTCTTG 120  
DB 255 AAGGAGAGAGCTCTGAGATCTCTTTGCTCTCCATATGAGGCAATCAGTGTG 314  
QY 121 GAATCTGGGCTGATTTGGTCTTTACTTGTGATTTGGCTGAGAAAGCTGCGCGATT 180  
DB 315 GAATCTGGGCTGATTTGGTCTTTACTTGTGATTTAGTCAAAAGGCTGCTGAACTC 374  
QY 181 ATTGCTCTGATTTCAATGAGAGTCTATCAAGAGAGATGAAGTAATCAATGGCGACTAC 240  
DB 375 ATTGCTCTGATTTCAATGAGAGTCTATCAAGAGAGATGAAGTAATCAATGGCGACTAC 434  
QY 241 AAAAAATCTGATTTATGCTGCTGATGCTGCTCTCCCACTCTCACTGTTTCCCAACCAT 300  
DB 435 AAGATGCTCAAGTTTATGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 494  
QY 301 TCATTTGATGATATTTCTCAATTTGTTTCTCAATTTCTGATTTCTTTCTGATGAAGAGTGA 360  
DB 495 TCTCTGACTGATTTCTTCTCAACTGGCTGCTCATGATCTTTCTGCAAAAGAGTGGAG 554  
QY 361 AATTTGCTTCAAGAGATTTGAAATGTTGAAAGCCAGGCTGATATTTCTTCTGAGAA 420  
DB 555 CTTTTGGCAGAGAGATGCTGGTGGATCAAGGTTGGAGATATTTCTTCTGCTGAA 614  
QY 421 TCTTTGTTTCAATCTGAGGATCAAAACCAAGAGATCCCAACCATCTACCGTGAA 480  
DB 615 TCTTTGCTTCAACCAATCAGGAGACAGTAAGCGGAAATCCAAACCCCACTCACTACCGTGAA 674  
QY 481 CTTAGGTTCTACACTAAGGCTTCAAGAGATGCTATTTGCAAGATGATCTGGAATCTCT 540  
DB 675 CCCCCTTCTTATCCAGGCTTTTCAAGAGTGTGAGATGCTGCTGCTGCTGCTGCTGCTG 734  
QY 541 TATGAGCTCTCCCTACTAGTCCAAATGATGAGCTTATGTCAGAAACAGAAAC 600  
DB 735 TTTGAGCTCTCTATGATCGGATGCAAGGCTGATGAGCTTATGTCAGAAACAGAAAT 794  
QY 601 CAGAACAGATTTAGTTGTTGCTGCAAAAGTGTATTTAAGATGATGAAGGCTTCCAG 660  
DB 795 CAGAAATCAGATTTGTTGATATGCGAAGAGTCAAGTCAAGAAATGACAGAGGCTTCAA 854  
QY 661 CGATTTCTGATATCTAGCAGTCAAGTGTATAGATTTCTCGGATATGAGCGTGTATTT 720  
DB 855 GCTTTCTTGCAATGCTCAATCAATCCAGTGGATCTTACGCTATGAGCGTGTCTTT 914  
QY 721 GCGCTGTTTATGTTAGCACTGAGGATATGAAACCAACCAAGATTTGTTGCTCAATGCTG 780  
DB 915 GCGCAAGGCTTTGTTGAGCACTGTTGAGCACTGAGCAACCAAGAAATTTGTTGAGAAATG 974  
QY 781 GACTTGAAGCTGCGCAGAGGCTCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 840  
DB 975 AATCTGAACCAAGCAGCAAGAGCTTTAGATGTTGGGCTGTTGTTGTTGTTGTTGTTGTTG 1034  
QY 841 TACATGGCGGAGACCTTTGATGTTGAGCTTTGTTGATTTGATCTCTCGTTAATGATTT 900  
DB 1035 TACATGGCTGAGAGTGTGATGTTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1094  
QY 901 TCTTTGCTCTGAGCGTTCTATTGCGCTTAAATGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 960  
DB 1095 TCTTTGCTCTGAGCGTTCTATTGCGCTTAAATGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1154  
QY 961 TGCACCAAGATTAATCACTACCTGATACTCTTTGATGCTATCTATAGCGCTGACACCATTT 1020  
DB 1155 TGCACCAAAACACTACCCAGATTAATCTGTTGATGCTATCTATAGCGCTGACACCATTT 1214  
QY 1021 CTGCAATTTAGGCAAGCGCTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1080  
DB 1215 CTGCAATTTAGGCAAGCAACACAGCGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1274  
QY 1081 GATAAGTTTCAATCAGTGTATCTGCAAGAAAGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140  
DB 1275 GATAAGTTTCAATCAGTGTATCTGCAAGAAAGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1334  
QY 1141 GCTTACATTAAGCAGAGGGGATATGATCTCCATGATGTAAGGAATATGGGCAGATGCTT 1200

DB 1335 GAGTACATCAACAGAGAGATATGATCTCCATGCTTCAAGCTTATGACAGATGCTA 1394  
QY 1201 AAGATGCTGGATTTGTTGATGCTTCTGCGAGGATAGACTGAGCAGTTTCATTCGAGTT 1260  
DB 1395 AAGAGCGTGGCTTCACTGATGATGATCGAGAGCCGCTACTGATCAGTTTATGCAAGTC 1454  
QY 1261 CTACGGAAGAACTAGAGACTGTTGAGAGGAAAGAGTGTTCATTTAGTGTCTCT 1320  
DB 1455 CTGAAACGTTGAAATTACACAGGCTGAGAGAGAAAGGAAATTCATCTCGACTTCTCC 1514  
QY 1321 GAGGAGATTTACATGACATTTGAGGCTTGCATGATTAAGTTGCGAGAGACTGCCAAG 1380  
DB 1515 AAGAGGATTTACGATGACATTTGAGGATGAGAGTCAAGCTGGAGGTTGTCATCG 1574  
QY 1381 GGTGAGCAACGATGGGCTGCTGTTGTTGCTGCAAGAGAGTGAAGAA 1426  
DB 1575 GATGAGCAAGAAATGGGAGCTTTTTCATGCGCAACAGAAATTAAGCAA 1620

RESULT 13  
AF197940  
LOCUS Arabidopsis thaliana SAM:phospho-ethanolamine N-methyltransferase  
DEFINITION Arabidopsis thaliana (thale cress)  
ACCESSION AF197940  
VERSION AF197940.1 GI:11890405  
KEYWORDS  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
1 (bases 1 to 1878)  
Bolognese, C.P. and McGraw, P.  
The isolation and characterization in yeast of a gene for  
Arabidopsis S-adenosylmethionine:phospho-ethanolamine  
N-methyltransferase  
Plant Physiol. 124 (4), 1800-1813 (2000)  
MEDLINE 20567827  
PUBMED 11115895  
REFERENCE 2 (bases 1 to 1878)  
Bolognese, C.P. and McGraw, P.  
Direct Submision  
TITLE Submitted (22-OCT-1999) Biology, Univ. of Maryland, BC, 1000  
JOURNAL Hilltop Circle, Baltimore, MD 21250, USA  
FEATURES  
location/Qualifiers  
1..1878  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/cultivar="Landesberg erecta"  
/db\_xref="taxon:3702"  
/map="64921 to 67861"  
/tissue\_type="stage two leaves"  
/clone\_lib="Minet W; Dufour ME and Lacroute F., Plant J  
(1992) 2:417-422"  
/dev\_stage="complete seedlings"  
/note="A. thaliana (L.) Heynh"  
1..1878  
/gene="NMT1"  
202..1677  
/gene="NMT1"  
/function="catalyzes the production of  
phosphatidylcholine"  
/codon\_start=1  
/product="SAM:phospho-ethanolamine N-methyltransferase"  
/protein\_id="AAG41121.1"  
/db\_xref="GI:11890406"  
/translation="MARSVEEEDIQNTWIEHSDLTVEAMMLDSRLDKEERPE  
VLSLPYBKGSLVLELGCAGIGRTGRLAOKAGELIALDFIDNVIKKESINGHYKVK  
EMCADVTSPDKITDGLDIIFSNWLLNLSKVELLAERMWGIVKGGYIFPESC  
FQSGDSKRSNPHYREPFYSKVFQECQTRDAAGNSFELSMICKICGATVVKKNK

QNCQWTKVQKSSSENRGQRFQFLDNVOYKSGILRYRVPQGFVSTGGJLTKKFVE  
 KMLKPKQKVLVDGCGIGGGDFYMAEKFDVHVVDLSDVNMISFALERAIGLSCSVF  
 EVADCTKYPDPSFVITVSDRTILHIQDPALFPTFKMLKPKQKVLISDYCRSPK  
 PSAPSEYIKORGYDLHDVQAYQKMLKADAGPTDVIAEDRTDQFMQVLKRELDVEREK  
 EKFTSDFSKDEYDDIVGGWKSLEKRCASDEQKGLFIANKN"

misc\_feature

361..660

/gene="NM1"  
 /note="Region: SAM-binding domain"

misc\_feature

361..387

/gene="NM1"  
 /note="Region: Motif I"

misc\_feature

427..438

/gene="NM1"  
 /note="Region: post-I"

misc\_feature

544..570

/gene="NM1"  
 /note="Region: Motif II"

misc\_feature

631..660

/gene="NM1"  
 /note="Region: Motif III"

misc\_feature

1048..1341

/gene="NM1"  
 /note="Region: SAM-binding domain"

misc\_feature

1048..1074

/gene="NM1"  
 /note="Region: Motif I"

misc\_feature

1114..1128

/gene="NM1"  
 /note="Region: post-I"

misc\_feature

1231..1254

/gene="NM1"  
 /note="Region: Motif II"

misc\_feature

1312..1341

/gene="NM1"  
 /note="Region: Motif III"

## ORIGIN

Query Match 52.1%; Score 834; DB 8; Length 1878;  
 Best Local Similarity 74.1%; Pred. No. 7.4e-195;  
 Matches 1056; Conservative 0; Mismatches 370; Indels 0; Gaps 0;

QY 1 CACACGTTGATTTAACCATGAGCTATGATGCTGCTGATTCACAGCTCTGATCTTGAC 60  
 DB 256 CATTCCGCTGATCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAC 315  
 QY 61 AAAGAAGAACGCTCTGAGATCTTTCAATGCTTCCGCTCTTTGAAGGAAATGCTCTTG 120  
 DB 316 AAGGAAGAACGCTCTGAGGACTCTCTTGTCTCTCCATATGAAGGCAATCACTGTTG 375  
 QY 121 GAACCTGGGCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
 DB 376 GAACCTGGGCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 435  
 QY 181 ATTGCTCTGATTTCAATGAGTGTCTATCAAGAAGATGAAGTAAATCAATGGGCACTAC 240  
 DB 436 ATTGCTCTGATTTCAATGAGTGTCTATCAAGAAGATGAAGTAAATCAATGGGCACTAC 495  
 QY 241 AAAAATGTCAGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
 DB 496 AAGAATGTCAGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 555  
 QY 301 TCATTGATGATATTTCTCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
 DB 556 TCTCTGATGATATTTCTCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 615  
 QY 361 AATTGCTGAAAGATTTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
 DB 616 CTTTGGCAGAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 675  
 QY 421 TCTGTTTCCATCAATCTGGGATCAAGACGCAAGACCAATCCCAACCACTACCGTGA 480  
 DB 676 TCTGTTTCCATCAATCTGGGATCAAGACGCAAGACCAATCCCAACCACTACCGTGA 735

QY 481 CTTAGGTTCTACACTAAGGCTTCAAGAGTGTCTATTTGCAAGATGATCTGAAATCTCT 540  
 DB 736 CCGCGTTCTATTTCAAGGCTTTTCAAGAGTGTCTGCTGCGGAGTGTCTGCTGAAATTTCA 795  
 QY 541 TATGAGCTCTCCCTACTTACTGCTCAAAATGATTTGAGAGCTTATGTCAAGAACCAAGAAAC 600  
 DB 796 TTTGAGCTCTCTATGATCGGATGCAATGCAATTTGAGAGCTTATGTCAAGAACCAAGAAAT 855  
 QY 601 CAGAACCAAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
 DB 856 CAGAACCAAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 915  
 QY 661 CGATTTCTGATACTACCCAGTACAAGTGTAAATGATCTGCGATATGAGGCTGTATTT 720  
 DB 916 CGTTTCTTGGACATGCTCAATACAAATCCAGTGGAACTCTAGCTATGAGGCTGTCTTT 975  
 QY 721 GGCCTGCTGATTTAGCACTGAGGATATGAACCAACCAAGAGCTTTGCTGCTGCTGCTGCT 780  
 DB 976 GGCCTGCTGATTTGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1035  
 QY 781 GACTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
 DB 1036 AATCTGAACCCAGACAGAAAGCTTTAGATTTGGGTGTGGCAATCTGCTGAGGCTGCTTC 1095  
 QY 841 TACATGGCGAGACCTTTGATTTGAGGTTTGGATTTGATCTCTCGTTTAAATGATTT 900  
 DB 1096 TACATGGCTGAGAGTTTGGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 1155  
 QY 901 TCCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960  
 DB 1156 TCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1215  
 QY 961 TGCACCAAGATAAACCCTGATTAATCTTTTGTATGCTCTATAGGCTGCTGACCACTTT 1020  
 DB 1216 TGCACCAACCAACACTACCAGATTAATCTTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 1275  
 QY 1021 CTGCATATTCAGGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
 DB 1276 CTGCATATTCAGGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1335  
 QY 1081 GGTAAAGTTCTATCTAGTATTAATCTGCAAGAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
 DB 1336 GGTAAAGTTCTATCTAGTATTAATCTGCAAGAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1395  
 QY 1141 GCTTACATTAAGCAGAGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
 DB 1396 GAGTACATCAACAGAGAGGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1455  
 QY 1201 AAAGATCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260  
 DB 1456 AAAGATCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1515  
 QY 1261 CTACGGAAGGAACTAGAGACTGTTGAGAAAGGAAAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320  
 DB 1516 CTGAAAAGCTGAAATAGACAGGCTGGAAGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAA 1575  
 QY 1321 GAGGAGATTAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380  
 DB 1576 AAAGAGATTAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1635  
 QY 1381 GGTGAGCAAGATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1426  
 DB 1636 GATGAGCAGAAATGGGAGCTTTTCTATCGCCCAACAGAAATTAAGCAA 1681

RESULT 14  
 AY091683

LOCUS AY091683 1476 bp mRNA linear PLN 13-APR-2002  
 DEFINITION Arabidopsis thaliana AT3g18000/ME85\_22 mRNA, complete cds.

ACCESSION AY091683  
 VERSION AY091683.1 GI:20147130

KEYWORDS FLI CDNA.  
 SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana  
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi  
 (bases 1 to 1476)  
 Cheuk,R., Chen,H., Kim,C.J., Meyers,M.C., Shinn,P., Banh,J.,  
 Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D.,  
 Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G.,  
 Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M.,  
 Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T.,  
 Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C.,  
 Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W.,  
 Theologis,A. and Ecker,J.R.  
 Arabidopsis ORF clones  
 Unpublished  
 2 (bases 1 to 1476)  
 Cheuk,R., Chen,H., Kim,C.J., Meyers,M.C., Shinn,P., Banh,J.,  
 Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D.,  
 Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G.,  
 Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M.,  
 Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T.,  
 Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C.,  
 Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W.,  
 Theologis,A. and Ecker,J.R.  
 Direct Submission  
 Submitted (21-MAR-2002) Salk Institute Genomic Analysis Laboratory  
 (SIGNAL), Plant Biology Laboratory, The Salk Institute for  
 Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,  
 USA  
 RIKEN Genomic Sciences Center (GSC) members carried out the  
 collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN  
 Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J.,  
 Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,  
 Hayashizaki,Y. and Shinozaki,K.  
 The Salk, Stanford, PBC (SSP) Consortium members carried out the  
 sequencing and annotation of the RAFL cDNAs : Cheuk,R., Chen,H.,  
 Kim,C.J., Meyers,M.C., Shinn,P., Banh,J., Bowser,L., Chan,M.M.,  
 Chang,E., Dale,J.M., Deng, J.M., Goldsmith,A.D., Jones,T.,  
 Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Miranda,M., Nguyen,M.,  
 Onodera,C.S., Palm,C.J., Quach,H.L., Southwick,A., Tang,C.C.,  
 Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Davis,R.W.,  
 Theologis,A., and Ecker,J.R.  
 Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to  
 this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)  
 contributed equally to this work as PIs.  
 location/Qualifiers  
 1..1476  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /chromosome="3"  
 /clone="U12246"  
 /ecotype="Columbia"  
 /note="This clone is in pUNI 51"  
 1..1476  
 /note="putative methyltransferase"  
 /codon\_start=1  
 /product="AT3g18000/MEB5.22"  
 /protein\_id="AA10282.1"  
 /db\_xref="GI:20147131"  
 /translation="MAASYEERDIOKNYWIHSADLTVEAMLDSDASLDKEERPE  
 VLSLPPEKSKVLEAGIGFTGELAQAGELALQFIDNVKQESINCHYKNV  
 FMCADVSPDLKTSGLSLDLSFNNLLATLSKEVELLAERHGVNKKYIFPREK  
 PQGSDKRSNPTHTPRFTSKVFCQCTRDAGNSPELSMGCKICGAYVNRKN  
 QNQLCPQKVVSENRGPFQFLDNVQYSSGLRYRFFVGQFVSTGLSTKEFVE  
 KXNLKPOKLVGCGGDFDYMAGKFDVHVVGIDLSVNMISFALERAGLSGVSE  
 EVADCTTKHYDPSFDVYSRDTLHIDQKLPALFTFKMLKPKLVLSIDYCRSPK  
 PSAESFYTHQGVLDHVDQYQMLKADGFTDVAERKTDQPMQVLKRELDREKEK  
 EKPTSDFSKEDYDDIVGWSKSLERCAESDEQXWGLFIANKN"

## ORIGIN

Query Match 52.0%; Score 833.2; DB 8; Length 1476;  
 Best Local Similarity 74.1%; Pred. No. 1,1e-194;  
 Matches 1054; Conservative 0; Mismatches 368; Indels 0; Gaps 0;  
 QY 1 CACACCGTTGATTTAACCAATTGAAGCTATGATGCTCGATTTCTCAAGCTTCTGATCTTGAC 60  
 DB 55 CATTCCCGCTGATCTGACTGTTGAAGCTATGATGCTTGACTCGAGAGCTTCTGATCTCGAC 114  
 QY 61 AAGAAGAACGCTCTCAGAGATTTCTTCAATGCTTCCGCTCTTGAAGGAAAAATGCTCTTG 120  
 DB 115 AAGGAGAGAGCTCTCAGAGTACTCTCTTGGCTCCCTCCATATGAAGGCAAAATCAGTGTG 174  
 QY 121 GAATCTGGGCTGATTTGCTGTTTACTGCTGAATTTGGCTGAGAAAGCTGGCCAGGTT 180  
 DB 175 GAATCTGGAGCTGATTTGCTGTTTCACTGCTGAATTTAGCTCAAAAGGCTGGTGAATC 234  
 QY 181 ATTGCTCTGATTTCAATGAGAGTCTATCAAGAGAGATGAAGTATCAATGAGGCACTAC 240  
 DB 235 ATTGCTCTTCACTTCATGATACGTTATCAAGAGAGATGAAGTATCAATGAGGCACTAC 294  
 QY 241 AAAAATCTCAAGTTTATGCTGCTGATGATCTTCTCCCACTCTCAGTCTTCCCAACCAT 300  
 DB 295 AAGAATGCTCAAGTTTATGCTGCTGATGATCTTCAATCCCTGACCTCAAGATCACTGATGA 354  
 QY 301 TCATTGATGATGATATTCTCCAAATGCTGATCTCTGATGATGATGATGATGATGATGAT 360  
 DB 355 TCTCTGATGATGATTTCTCCAAATGCTGATCTCTGATGATGATGATGATGATGATGAT 414  
 QY 361 AATTGCTGAAAGATTTGAAATGTTGAAGCAGGGGTTACATTTCTTCAGAGAA 420  
 DB 415 CTTTGGCAGAAAGGATGCTGCTGCTGATCAAGGTTGGAGGATACATTTCTTCGCTGAA 474  
 QY 421 TCTTGTTCATCAATCTGGGATCAAAACCAAGCAATCCCAACCACTACCTCCGTTGAA 480  
 DB 475 TCTTGTCTCCCAATCAAGGAGACAGTGAAGCGGAAATCCACCCCACTCACTACCGTGA 534  
 QY 481 CTTAGGTTCTACATAAGGCTTCAAGAGTGTCTATTTGCAAGATGGAATCGGAAACTCT 540  
 DB 535 CCGGCTTCTATTTCAAGAGTCTTTCAGAGTGTGACAGCTCGGGATGCTGCTGGAAATTC 594  
 QY 541 TATGAGCTCTCCCTACTACTAGCTCAAAATGATTTGAGAGCTTATGTCAGAAACCAAGAAAC 600  
 DB 595 TTTGAGCTCTCTATGATCGGATCAAGTGCATTTGAGCTTATGTCAGAAACCAAGAAAT 654  
 QY 601 CAGACAGATGATGTTGCTGCAAAAGTGTCTTCAAGATGATGAAGGTTCCAG 660  
 DB 655 CAGATGATGATGTTGCTGCAAAAGTGTCTTCAAGATGATGAAGGTTCCAG 714  
 QY 661 CGATTTCTGAGTACTAGCCAGTACAAAGTGAATGATGATGATGATGATGATGATGATGAT 720  
 DB 715 CGTTTCTTGGACAAATGTTCCAAATGATGATGATGATGATGATGATGATGATGATGATGAT 774  
 QY 721 GGCTCTGTTATGTTAGCACTGAGAGATGAAACCAACCAAGAGTGTGTTGTTCAATGCTG 780  
 DB 775 GGCCAAAGGTTTGTGAGCACTGCTGGAATGAGACAAACCAAGAAATTTGTCGAGAAATG 834  
 QY 781 GACTTGAAGCTCGCCAGAGAGTCTGAGATGTTGGTGTGGAATTTGGTGGAGGAGTCACTT 840  
 DB 835 AATCTGAACCCAGACAGAAAGTCTTAGATGTTGGGTTGAGCAATTTGGTGGAGGAGTCACT 894  
 QY 841 TACATGGCGGAGACTTTGATGTTGAGGTTGTTGGAATTTGATCTCTCCGTTAAATATGAT 900  
 DB 895 TACATGGCTGAGAGATTTGATGTTTCAAGTGTGTTGATGATGATGATGATGATGATGAT 954  
 QY 901 TCTTTGCTTGAAGCTTCTATGAGGCTTAATGCTGCTGATGATGATGATGATGATGATGAT 960  
 DB 955 TCTTTGCAATTTGAAGCTGCTTATGAGCACTGAGCTGCTGCTGCTGCTGCTGCTGCTGAT 1014  
 QY 961 TGCAACCAAGAACTACCTCGATTAATCTTTTGTGATGCTATCTATAGCGGTGACCACTT 1020  
 DB 1015 TGCAACCAAAACACTACCAAGCAATTTGTTAGTGTCAATTTACAGCGGTGACCACTATT 1074  
 QY 1021 CTGATATTTCAGGACAGCGCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1080







Db 870 CAGAAATCAGATTGTTGGATATGCGCAGAGGTGAGCTCAGAAATGACAGAGGCTTCCAA 929  
Qy 661 CGATTTCCTGGATCTAGCAGTCAAGGTGTAATAGCAATCTCGGATATGAGCGTGTATTT 720  
Db 930 CGTTTCCTGGCAATGTCCAAATCCAAATCCAGTGGATCTAGGCTATGAGCGTGTCTTT 989  
Qy 721 GGGCTGGGTATGTTAGCAGTGGAGGATATGAACCAACCAAGAGTTTGTGCAATGCTG 780  
Db 990 GGGCAAGGTTTGTGAGCACTGGTGGACTTGAGCAACCAAGAAATTTGTGAGAAATG 1049  
Qy 781 GACTTGAAGCTGGCCAGAGGCTCTGGATGTTGGTGTGGAAATGGTGGAGTGCATTT 840  
Db 1050 AATCTGAACACCAAGCAGACAGAAAGTCTTAGATGTTGGGTGGCATGCGTGGAGGTGACTTC 1109  
Qy 841 TACATGGCGGAGACCTTTGATGTTGAGTGTGGATTTGGATCTCTCCGTTAATATGAT 900  
Db 1110 TACNTGGCTGAGAGTTGATGTTGACGTTGTTGGTATGCAATCTTCTGTCACATGATC 1169  
Qy 901 TCCTTGGCCCTGAGCGTTCTATTTGGGCTTAAATGTGCTGTTGAGTTGAGTAGCAGAT 960  
Db 1170 TCTTTGGCATTGGAAACGCTGCTATTTGGACTCAGCTGCTCGGTTGGGTTGGGTTGCTGAT 1229  
Qy 961 TGCACCAAGATAAACTACCTGATAACTCTTTTGTGTCATCTATAGCCGTCACACCACTT 1020  
Db 1230 TGCACCAAAACACTACCCAGATAATTCGTTGATGTCATTTACAGCCGTCACACTAT 1289  
Qy 1021 CTGCATATTCAGGACAGCCCTGCTGTTTGTAGTCTTCTCAAAATGGTTGAAGCCAGGA 1080  
Db 1290 CTGCATCTCAAGACAAACCAAGCCTGTTTGTAGGACTTTCTTCAAAATGGCTTAAACCGGA 1349  
Qy 1081 GGTAAAGTTCTAATCAGTGTATTACTGCAAGAAAGCTGCTCCACCTCAGCTGAAATTCGCC 1140  
Db 1350 GGTAAAGTTCTCATCAGCGACTACTGTAGAAGCCCAAACTCCATCTGCTGATTTCA 1409  
Qy 1141 GCTTACATTAAGCAGAGGGGATATGATCTCCATGATGTAAAGGAATATGGCAGATGCTT 1200  
Db 1410 GAGTACATCAACAGAGAGGATATGATCTCCTGACGTTTCAAGCTTATGACAGATGCTA 1469  
Qy 1201 AAGATGCTGGATTTGTTGATGTTCTTGGCGAGGATAGAACTGAGCAGTTCAATCGAGTT 1260  
Db 1470 AAGAGCGTGGCTTCACTGATGTGATCGCAGAGGACCGTACTGATCAGTTTATGCAAGTC 1529  
Qy 1261 CTACGAGGAACTAGAGACTGTTGAGAAAGAAAGGATGTTCAATAGTATTTCTCT 1320  
Db 1530 CTGAAAGTGAATTAGACAGGCTGGAGAAAGAAAGAAATTCATCTCCGACTTCTCC 1589  
Qy 1321 GAGGAGTTTCAATGACATTTGTTGGAGGTTGGATGATGATGTTGGAGGACTGCCAAG 1380  
Db 1590 AAGAGGATTACGATGACATTTGTTGGAGGATGGAAGTCAAAGCTGGAGAGGTTGATCG 1649  
Qy 1381 GGTGAGCAACGATGGGCTCTGCTGTTGCCAAGAAAGTGAAGAA 1426  
Db 1650 GATGAGCAAAATGGGACTTTTCATCGCCAAACAAAGAAATTAAGCAA 1695

Search completed: August 2, 2004, 07:03:32  
Job time : 6486 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2004, 08:20:03 ; Search time 6481 Seconds  
(without alignments)  
10713.708 Million cell updates/sec

Title: US-10-031-331b-39  
Perfect score: 1602  
Sequence: 1 cacacgttgatttaaccat.....aaaaaaaaaaaaaaaaaaaa 1602

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_hg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vl.\*

15: gb\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vl.\*

30: em\_hg\_hum.\*

31: em\_hg\_inv.\*

32: em\_hg\_other.\*

33: em\_hg\_mus.\*

34: em\_hg\_pln.\*

35: em\_hg\_rod.\*

36: em\_hg\_nam.\*

37: em\_hg\_vrt.\*

38: em\_sy.\*

39: em\_hgo\_hum.\*

40: em\_hgo\_mus.\*

41: em\_hgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1602	100.0	1602	6	BD084055	BD084055 Environme
2	1602	100.0	1602	6	BD093370	BD093370 Environme
3	1602	100.0	1602	8	AB080186	AB080186 Sueda ja
4	35	2.2	1459	8	AF133811	AF133811 Arabidops
5	35	2.2	1656	8	AF248454	AF248454 Arabidops
6	35	2.2	1660	8	AY063866	AY063866 Arabidops
7	35	2.2	95235	8	AC020889	AC020889 Genomic e
8	32	2.0	1120	6	AX250553	AX250553 Sequence
9	32	2.0	2235	6	AX250551	AX250551 Sequence
10	32	2.0	2252	8	AF237633	AF237633 Spinacia
11	31	1.9	1347	5	BC049504	BC049504 Danio rer
12	31	1.9	138116	2	BX324004	BX324004 Danio rer
13	31	1.9	168179	2	BX293550	BX293550 Danio rer
14	31	1.9	172587	5	BX005125	BX005125 Zebrafish
15	31	1.9	175745	2	BX530060	BX530060 Danio rer
16	31	1.9	184574	5	BX294666	BX294666 Zebrafish
17	31	1.9	187570	2	BX649276	BX649276 Danio rer
18	31	1.9	226787	2	BX004886	BX004886 Danio rer
19	31	1.9	247330	2	BX530409	BX530409 Danio rer
20	31	1.9	251370	5	AL929533	AL929533 Zebrafish
21	30	1.9	1875	9	BC053645	BC053645 Homo sapi
22	30	1.9	2600	9	BC055010	BC055010 Homo sapi
23	30	1.9	105045	9	AC117494	AC117494 Homo sapi
24	30	1.9	129782	5	BX120005	BX120005 Zebrafish
25	30	1.9	180305	5	BX004768	BX004768 Zebrafish
26	30	1.9	197549	5	AL935335	AL935335 Zebrafish
27	30	1.9	347050	3	PFA929351	PFA929351 Plasmodiu
28	29	1.8	786	3	AB080223	AB080223 Samia cyn
29	29	1.8	1016	8	AY349618	AY349618 Arabidops
30	29	1.8	1102	9	BC016308	BC016308 Homo sapi
31	29	1.8	1429	8	AF508150	AF508150 Rheum tat
32	29	1.8	1449	9	BC005331	BC005331 Homo sapi
33	29	1.8	1734	9	AK000538	AK000538 Homo sapi
34	29	1.8	1833	9	BC000954	BC000954 Homo sapi
35	29	1.8	1870	8	AB090883	AB090883 Aster tri
36	29	1.8	1891	9	HSM807589	HSM807589 Homo sapi
37	29	1.8	2742	9	HSM801378	HSM801378 Homo sapi
38	29	1.8	3518	9	BC007647	BC007647 Homo sapi
39	29	1.8	6514	5	BC063203	BC063203 Silurana
40	29	1.8	61793	2	AC087703	AC087703 Homo sapi
41	29	1.8	71284	8	F21B23	F21B23 Arabidops
42	29	1.8	79389	10	AL607127	AL607127 Mouse DNA
43	29	1.8	83141	9	AL592067	AL592067 Human DNA
44	29	1.8	85702	8	AC020665	AC020665 Arabidops
45	29	1.8	110000	2	PFMAU13_02	Continuation (3 of

# ALIGNMENTS

RESULT 1	LOCUS	BD084055	1602 bp	DNA	linear	PAT 27-AUG-2002
DEFINITION	BD084055	Environmental stress-tolerant gene.				
ACCESSION	BD084055	Environmental stress-tolerant gene.				
VERSION	BD084055.1	GI:22629665				
KEYWORDS	JF 2001333784-A/20.	unidentified				
SOURCE	unidentified	unclassified.				
ORGANISM	unclassified.	unclassified.				
REFERENCE	1 (bases 1 to 1602)					
AUTHORS	Yamada,A., Ozeki,Y. and Saito,T.					
TITLE	Environmental stress-tolerant gene					
JOURNAL	Patent: JP 2001333784-A 20 04-DEC-2001;					
	JAPAN SCIENCE AND TECHNOLOGY CORP					

COMMENT	OS	Sueada japonica	781	GA	CTTTGAAGCCTGCGCAAGAGGTCCTGGATGTTGGTTGGGAATTTGGTGGAGGTGACATT	840
	PN	JP 2001333784-A/20	781	GA	CTTTGAAGCCTGCGCAAGAGGTCCTGGATGTTGGTTGGGAATTTGGTGGAGGTGACATT	840
	PD	04-DEC-2001				
	PP	19-JUL-2000 JP 2000219649				
	PI	AKIYO YAMADA, YOSHIHIRO OZEKI, TAKBO SAITO	841	TAC	ATGCGCGGAGAGCCTTTGATGTTGGAGTTGTTGGATTTGATCTCTCCGTTAATATGATT	900
	PC	C12N15/09, A01H5/00, C07K14/415, C07K16/16, C12N5/10, C12P21/02//	841	TAC	ATGCGCGGAGAGCCTTTGATGTTGGAGTTGTTGGATTTGATCTCTCCGTTAATATGATT	900
	PC	C12N1/21, C12N15/00, C12N5/00				
	PC	C12P21/08, C12N15/00, C12N5/00				
	CC	Environmental stress-tolerant gene	901	TC	CTTTGGCCTTGAGCGTTCTATTGGCTTAAATGCTGTTGAGTTTGAAGGTAGCAGAT	960
	FH	Key	901	TC	CTTTGGCCTTGAGCGTTCTATTGGCTTAAATGCTGTTGAGTTTGAAGGTAGCAGAT	960
	FT	CDS				
FEATURES		Location/Qualifiers				
source		1..1602				
		/organism="unidentified"				
		/mol_type="genomic DNA"				
		/db_xref="taxon:32644"				
ORIGIN						
	Query Match	100.0%; Score 1602; DB 6; Length 1602;				
	Best Local Similarity	100.0%; Pred. No. 0;				
	Matches 1602; Conservative	0; Mismatches 0; Indels 0; Gaps 0;				
Qy	1	CACACCGTTGATTAACCAATGAGCTATGATGCTCGATCTCAAGCTTCTGATCTGAC	60			
Db	1	CACACCGTTGATTAACCAATGAGCTATGATGCTCGATCTCAAGCTTCTGATCTGAC	60			
Qy	61	AAAGAGAGCGTCTGAGATCTTTCAATGCTTCGGCTCTTGAAGGAAATGCTCTTG	120			
Db	61	AAAGAGAGCGTCTGAGATCTTTCAATGCTTCGGCTCTTGAAGGAAATGCTCTTG	120			
Qy	121	GAACTTGGGCTGATGTTGGTTCCTTTACTGTTGATGATGCTGAGAAAGCTGCGAGTT	180			
Db	121	GAACTTGGGCTGATGTTGGTTCCTTTACTGTTGATGATGCTGAGAAAGCTGCGAGTT	180			
Qy	181	ATTGCTCTGATTTCAATGAGAGTGTCTATCAAGAGAGATGAATCAATGGGCACTAC	240			
Db	181	ATTGCTCTGATTTCAATGAGAGTGTCTATCAAGAGAGATGAATCAATGGGCACTAC	240			
Qy	241	AAATATGTCAGTTTATGCTGATGCTGATGCTTCCACTCTCAGTTTCCACCAAT	300			
Db	241	AAATATGTCAGTTTATGCTGATGCTTCCACTCTCAGTTTCCACCAAT	300			
Qy	301	TCATTGATGATGATATCTCCAAATGGTTACTCATGTATCTTTCTGATGAAGAGGTGAA	360			
Db	301	TCATTGATGATGATATCTCCAAATGGTTACTCATGTATCTTTCTGATGAAGAGGTGAA	360			
Qy	361	AATTTGGTTCAAGAGATGTTGAATGCTTGAAGCAGGGGTTACATTTCTTCAGAGAA	420			
Db	361	AATTTGGTTCAAGAGATGTTGAATGCTTGAAGCAGGGGTTACATTTCTTCAGAGAA	420			
Qy	421	TCCTTTTCCATCAATCTGGGATCAAAACGCAAAAGCAATCCCACTACCGTGAA	480			
Db	421	TCCTTTTCCATCAATCTGGGATCAAAACGCAAAAGCAATCCCACTACCGTGAA	480			
Qy	481	CTTAGTTCTACACTAAGGCTCTCAAGAGTGTCTATTTGCAAGATGGAATCTGAAACTCT	540			
Db	481	CTTAGTTCTACACTAAGGCTCTCAAGAGTGTCTATTTGCAAGATGGAATCTGAAACTCT	540			
Qy	541	TATGAGCTCTCCCTACTAGCTCAATGATGATGAGCTTATGTCAGAAACAGAAAC	600			
Db	541	TATGAGCTCTCCCTACTAGCTCAATGATGATGAGCTTATGTCAGAAACAGAAAC	600			
Qy	601	CAGAACCCAGATTTAGTTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	660			
Db	601	CAGAACCCAGATTTAGTTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	660			
Qy	661	CGATTTCTGGATCTAGCCAGTCAAGTGTATGATGATTTCTGCGATATGAGCTGATTT	720			
Db	661	CGATTTCTGGATCTAGCCAGTCAAGTGTATGATGATTTCTGCGATATGAGCTGATTT	720			
Qy	721	GGCCCTGTTATGTTAGCTGAGGATATGAAACCAACCAAGAGTTTGTGTCAATGCTG	780			
Db	721	GGCCCTGTTATGTTAGCTGAGGATATGAAACCAACCAAGAGTTTGTGTCAATGCTG	780			

RESULT 2

BD093370

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

1602 bp

DNA

linear

PAT 27-AUG-2002

Environmental stress-tolerant gene.

BD093370.1

GI:22638958

WO 0106006-A/20.

unidentified

unclassified.

1 (bases 1 to 1602)

Yamada, A., Ozeki, Y. and Saito, T.

Environmental stress-tolerant gene

Patent: WO 0106006-A 20 25-JAN-2001.

JAPAN SCIENCE AND TECHNOLOGY CORP, AKIYO YAMADA, YOSHIHIRO OZEKI,

COMMENT  
 TAKEO SAITO  
 OS Sueaeta japonica  
 PN WO 0106006-A/20  
 PD 25-JAN-2001  
 PF 19-JUL-2000 WO 2000JP004862  
 PR 19-JUL-1999 JP 99P 235910.24-MAR-2000 JP OOP 085377 PI  
 PC AKIYO YAMADA, YOSHIIHRO OZeki, TAKEO SAITO  
 CI 201/58, C12N15/29, C07K14/415, C07K16/16, C12P21/02, A01H5/00 CC  
 FH Key Location/Qualifiers  
 FT CDS (1). (1419).  
 FEATURES  
 source Location/Qualifiers  
 1. 1602  
 /organism="unidentified"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32644"  
 ORIGIN  
 Query Match 100.0%; Score 1602; DB 6; Length 1602;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1602; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CACACCGTTCATTAAACCAATGAAGCTATGATGCTCGATTTCTCAAGCTTCTGATCTTGAC 60  
 DB 1 CACACCGTTCATTAAACCAATGAAGCTATGATGCTCGATTTCTCAAGCTTCTGATCTTGAC 60  
 QY 61 AAAGAAGAAGCTCCTGAGATTTCTTCAATGCTTCGGCTCTTGAAGGAAGTCCCTTGG 120  
 DB 61 AAAGAAGAAGCTCCTGAGATTTCTTCAATGCTTCGGCTCTTGAAGGAAGTCCCTTGG 120  
 QY 121 GAACTTGGGGCTGGATATGCTGCTTTTACTGGTGAATGCTGAGAAAGCTGCCAGGTT 180  
 DB 121 GAACTTGGGGCTGGATATGCTGCTTTTACTGGTGAATGCTGAGAAAGCTGCCAGGTT 180  
 QY 181 ATTGCTCTGGATTTCAATTGAGAGTGTCTATCAAGAAGAAATGAAGTAATCAATGGGCACTAC 240  
 DB 181 ATTGCTCTGGATTTCAATTGAGAGTGTCTATCAAGAAGAAATGAAGTAATCAATGGGCACTAC 240  
 QY 241 AAAAATGTCAGCTTATGCTGCTGATGCTGATTTCTCCCACTCTCAGATTTCCACCAAT 300  
 DB 241 AAAAATGTCAGCTTATGCTGCTGATGCTGATTTCTCCCACTCTCAGATTTCCACCAAT 300  
 QY 301 TCATTGGATGTGATATTTCCAAATGGTTTGAAGCCAGGGGTTTACATTTCTTCAGAGAA 360  
 DB 301 TCATTGGATGTGATATTTCCAAATGGTTTGAAGCCAGGGGTTTACATTTCTTCAGAGAA 360  
 QY 361 AATTTGGTTGAAGAAGATGTTGAATGGTTGAAGCCAGGGGTTTACATTTCTTCAGAGAA 420  
 DB 361 AATTTGGTTGAAGAAGATGTTGAATGGTTGAAGCCAGGGGTTTACATTTCTTCAGAGAA 420  
 QY 421 TCTTGTTCATCAATCTGGGGATCAAAAGCAAGCAATCCCACTACCGTGAA 480  
 DB 421 TCTTGTTCATCAATCTGGGGATCAAAAGCAAGCAATCCCACTACCGTGAA 480  
 QY 481 CTTAGGTTCTACATPAAGGCTCTCAAGAGTGTCTATTTGCAAGATGGATCTGGAATCT 540  
 DB 481 CTTAGGTTCTACATPAAGGCTCTCAAGAGTGTCTATTTGCAAGATGGATCTGGAATCT 540  
 QY 541 TATGAGCTCTCCCTACTTAGCTGCAATGATTTGGAGCTTATCTCAGAACTCT 600  
 DB 541 TATGAGCTCTCCCTACTTAGCTGCAATGATTTGGAGCTTATCTCAGAACTCT 600  
 QY 601 CAGAACCAAGATAGTTGGTTGTGGCAAAAAGTTGATTTAAAGGATGATTAAGGGGTTCCAG 660  
 DB 601 CAGAACCAAGATAGTTGGTTGTGGCAAAAAGTTGATTTAAAGGATGATTAAGGGGTTCCAG 660  
 QY 661 CGATTTCTGGATCTAGCCAGTACAGTGTAAATGATCTCTCGATATGAGCTGATTT 720  
 DB 661 CGATTTCTGGATCTAGCCAGTACAGTGTAAATGATCTCTCGATATGAGCTGATTT 720  
 QY 721 GGGCCCTGTTATGTTAGCACTGGAGGATFAGAAACGACCAAGAGTGTGTCAATGCTG 780  
 DB 721 GGGCCCTGTTATGTTAGCACTGGAGGATFAGAAACGACCAAGAGTGTGTCAATGCTG 780

QY 781 GACTTCGAAGCCTGGCCAGAAAGTCTCGATGTTGGTGTGGAAATGGTGAGGTGACTTT 840  
 DB 781 GACTTCGAAGCCTGGCCAGAAAGTCTCGATGTTGGTGTGGAAATGGTGAGGTGACTTT 840  
 QY 841 TACATGGCGGAGACCTTTGATGTTGAGGTGTTGAGATTTGATCTCTCCGTTAATATGAT 900  
 DB 841 TACATGGCGGAGACCTTTGATGTTGAGGTGTTGAGATTTGATCTCTCCGTTAATATGAT 900  
 QY 901 TCTTTTGGCCCTTGAGCGCTTATTTGGGCTTAAATGTCCTGTTGAGTTTGGAGTAGCAGAT 960  
 DB 901 TCTTTTGGCCCTTGAGCGCTTATTTGGGCTTAAATGTCCTGTTGAGTTTGGAGTAGCAGAT 960  
 QY 961 TGCACCAAGATATAATCTACCTCGATTAATCTTTTGAATGTCATCTATAGCGTGCACACAT 1020  
 DB 961 TGCACCAAGATATAATCTACCTCGATTAATCTTTTGAATGTCATCTATAGCGTGCACACAT 1020  
 QY 1021 CTGCATATTTCAAGCAAGCTCTGCGTTGTTTAGATCTTTTACAAATGGTTGAAGCCAGCA 1080  
 DB 1021 CTGCATATTTCAAGCAAGCTCTGCGTTGTTTAGATCTTTTACAAATGGTTGAAGCCAGCA 1080  
 QY 1081 GGTAAAGTTCTAATCAGTGTATTACTGCAAGAAAGCTGCTCCACCTCACTGAAATTCGCC 1140  
 DB 1081 GGTAAAGTTCTAATCAGTGTATTACTGCAAGAAAGCTGCTCCACCTCACTGAAATTCGCC 1140  
 QY 1141 GCTTACATTTAAGCAGAGGGGATATGATCTCCATGATGTAAGGAATATGCGCAGATGCTT 1200  
 DB 1141 GCTTACATTTAAGCAGAGGGGATATGATCTCCATGATGTAAGGAATATGCGCAGATGCTT 1200  
 QY 1201 AAAGATGCTGGAATTTGTTGATGTTCTTCCGAGGATAGAACTCAGCAGTTTCAATCGAGTT 1260  
 DB 1201 AAAGATGCTGGAATTTGTTGATGTTCTTCCGAGGATAGAACTCAGCAGTTTCAATCGAGTT 1260  
 QY 1261 CTACCGAAGGACTAGACGCTTTGAGAGAAAGATGTTGTTCAATAGTATTTCTCT 1320  
 DB 1261 CTACCGAAGGACTAGACGCTTTGAGAGAAAGATGTTGTTCAATAGTATTTCTCT 1320  
 QY 1321 GAGGAGGATTAACATGACATTTGTTGAGGTTGGAATGATTAAGTTGCGAGGACTGCCAAG 1380  
 DB 1321 GAGGAGGATTAACATGACATTTGTTGAGGTTGGAATGATTAAGTTGCGAGGACTGCCAAG 1380  
 QY 1381 GGTGAGCAACGATGGGCTGCTGTTGCGCAAGAGAGTGAAGATCAGTTGCCGACT 1440  
 DB 1381 GGTGAGCAACGATGGGCTGCTGTTGCGCAAGAGAGTGAAGATCAGTTGCCGACT 1440  
 QY 1441 GGCATCTGCAATTTCTAGTATTAATCTCAGATTTTCAATGTTTCAATGTTTCTTACATGT 1500  
 DB 1441 GGCATCTGCAATTTCTAGTATTAATCTCAGATTTTCAATGTTTCAATGTTTCTTACATGT 1500  
 QY 1501 AAAATTTGCCAATAAGTTGCAATTTCCAGACTGTAAGATGATTAATCATATTTTATCTTTT 1560  
 DB 1501 AAAATTTGCCAATAAGTTGCAATTTCCAGACTGTAAGATGATTAATCATATTTTATCTTTT 1560  
 QY 1561 AATTATCATGATTTATGCAAAAAAAGAAAAAAGAAAAA 1602  
 DB 1561 AATTATCATGATTTATGCAAAAAAAGAAAAAAGAAAAA 1602

RESULT 3  
 AB080186  
 LOCUS  
 DEFINITION  
 Sueaeta japonica PEAMT mRNA for phosphoethanolamine  
 N-methyltransferase, complete cds.  
 ACCESSION  
 AB080186  
 VERSION  
 AB080186.1 GI:28436073  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Sueaeta japonica  
 Sueaeta japonica  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 Caryophyllales; Amaranthaceae; Sueaeta.  
 REFERENCE  
 1 Yamada, A., Nozawa, G.T., Tanimoto, S. and Ozeki, Y.  
 Glycinebetaine synthesis in Sueaeta japonica

Unpublished		541		TATGAGCTCTCCCTA	QY	541		TATGAGCTCTCCCTA	QY
2 (bases 1 to 1779)		718		TATGAGCTCTCCCTA	Db	718		TATGAGCTCTCCCTA	Db
Yamada, A., Nozawa, G.T. and Ozeki, Y.		601		CAGAACACAGATTA	QY	601		CAGAACACAGATTA	QY
Direct Submission		778		CAGAACACAGATTA	Db	778		CAGAACACAGATTA	Db
Submitted (19-FEB-2002) Akiyo Yamada, Tokyo University of		661		CGATTTCTGGATTA	QY	661		CGATTTCTGGATTA	QY
Agriculture and Technology, Department of Biotechnology; Naka-cho		838		CGATTTCTGGATTA	Db	838		CGATTTCTGGATTA	Db
2-24-16, Koganei, Tokyo 184-8588, Japan		721		GGCCCTGGTATTA	QY	721		GGCCCTGGTATTA	QY
(E-mail: yamadenec.tuat.ac.jp, Tel: 81-423-88-7239,		898		GGCCCTGGTATTA	Db	898		GGCCCTGGTATTA	Db
Fax: 81-423-88-7239)		781		GACTTGAGGCTGG	QY	781		GACTTGAGGCTGG	QY
Location/Qualifiers		958		GACTTGAGGCTGG	Db	958		GACTTGAGGCTGG	Db
1. .1779		841		TATATGGGAGAC	QY	841		TATATGGGAGAC	QY
/organism="Suaeda japonica"		1018		TATATGGGAGAC	Db	1018		TATATGGGAGAC	Db
/db_xref="taxon:90346"		901		TCCTTTGGCCCTT	QY	901		TCCTTTGGCCCTT	QY
/country="Japan: Ariake Sea"		1078		TCCTTTGGCCCTT	Db	1078		TCCTTTGGCCCTT	Db
1. .1779		961		TGACCAAGATTA	QY	961		TGACCAAGATTA	QY
/gene="PEAMT"		1138		TGACCAAGATTA	Db	1138		TGACCAAGATTA	Db
115. .1599		1021		CTGCATATTCAG	QY	1021		CTGCATATTCAG	QY
/codon_start=1		1198		CTGCATATTCAG	Db	1198		CTGCATATTCAG	Db
/product="phosphoethanolamine N-methyltransferase"		1081		GCTAATAGTCTA	QY	1081		GCTAATAGTCTA	QY
/protein_id="BAC57432.1"		1258		GCTAATAGTCTA	Db	1258		GCTAATAGTCTA	Db
/db_xref="GI:28436074"		1141		GCTTACATTAAG	QY	1141		GCTTACATTAAG	QY
/translation="MAASGMEIOMERDVFVKYVHTVLTITIAWMLDSQASDLKKEE		1318		GCTTACATTAAG	Db	1318		GCTTACATTAAG	Db
RPELSMLPPLGKCLRLGAGIGRFTGELAEKAGOVIALDFIESAKKNEVINGHYK		1201		AAAGATGCTGGA	QY	1201		AAAGATGCTGGA	QY
NYKFCADVTSPSPSPHSDVLFNSMLWLSDEVNLVERMLKLPKGYIFFR		1378		AAAGATGCTGGA	Db	1378		AAAGATGCTGGA	Db
ESCFQSGDHKKSNPHYRPRFTYAPKECHLODGSNSVELSLSCKICGATVRN		1261		CTACCGNAGNACT	QY	1261		CTACCGNAGNACT	QY
KNQNOISLWQKVDKODKGFQFLDTSQTKNSILRYRFRVPGVSTGVBHTKE		1438		CTACCGNAGNACT	Db	1438		CTACCGNAGNACT	Db
FVSMIDLKPGQVLDVGGIGGGDFYMAETFDVEVGFSDLSVNMISFALERSIGLKA		1321		GAGGAGATTPAC	QY	1321		GAGGAGATTPAC	QY
VFVEADCTKNYPDPSFDVTSRDITLHIDPKPFRFYLKPGKVLISDYCKK		1498		GAGGAGATTPAC	Db	1498		GAGGAGATTPAC	Db
AGPPSPFAATYKQRYDLHDKVYKQKMDAGFVDVLAEDTQPIRVLRKELETVE		1381		GGTGAGCAACG	QY	1381		GGTGAGCAACG	QY
KEKDVFISDSSEYNDIVGGNDKLRRTAKGEQWGLFVAKKK"		1558		GGTGAGCAACG	Db	1558		GGTGAGCAACG	Db
ORIGIN		1441		GGGACTGTCCAT	QY	1441		GGGACTGTCCAT	QY
Query Match 100.0%; Score 1602; DB 8; Length 1779;		1618		GGGACTGTCCAT	Db	1618		GGGACTGTCCAT	Db
Best Local Similarity 100.0%; Pred. No. 0;		1501		AAAAATTCGCA	QY	1501		AAAAATTCGCA	QY
Matches 1602; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		1678		AAAAATTCGCA	Db	1678		AAAAATTCGCA	Db
QY 1 CACACCGTTGATTAACCAATGAAGCTATGATGCTGATGCTCAAGCTTCGATCTTGAC 60		1561		AAATTAATCAT	QY	1561		AAATTAATCAT	QY
Db 178 CACACCGTTGATTAACCAATGAAGCTATGATGCTGATGCTCAAGCTTCGATCTTGAC 237		1738		AAATTAATCAT	Db	1738		AAATTAATCAT	Db
QY 61 AAGAGAAAGCTCTCAGATTCCTTCAATGCTTCGGCTCTTGAAGGAAATGCTCTTG 120									
Db 238 AAGAGAAAGCTCTCAGATTCCTTCAATGCTTCGGCTCTTGAAGGAAATGCTCTTG 297									
QY 121 GAATCTGGGCTGGATTCGCTGTTTACTGCTGATTCGCTGATTCGCTGATTCGCTGATTC 180									
Db 298 GAATCTGGGCTGGATTCGCTGTTTACTGCTGATTCGCTGATTCGCTGATTCGCTGATTC 357									
QY 181 ATTGCTCTGGATTCATTCAGAGTGTCTATCAAGAGAAATGAAGTAATCAATGGGCACTAC 240									
Db 358 ATTGCTCTGGATTCATTCAGAGTGTCTATCAAGAGAAATGAAGTAATCAATGGGCACTAC 417									
QY 241 AAAAATCTCAAGTTTATGCTGATGCTGATTCCTCCACTCTCAGTTTCCACACCAT 300									
Db 418 AAAAATCTCAAGTTTATGCTGATGCTGATTCCTCCACTCTCAGTTTCCACACCAT 477									
QY 301 TCATGATGATGATTCCTCAATGGTTACTCATGATTCATGATTCCTTCATGATGATGATG 360									
Db 478 TCATGATGATGATTCCTCAATGGTTACTCATGATTCATGATTCCTTCATGATGATGATG 537									
QY 361 AAATTTGTTGAAAGAAATGTTGAAATGTTGAAAGCCAGGGGTTACATTTCTTCAGAGAA 420									
Db 538 AAATTTGTTGAAAGAAATGTTGAAATGTTGAAAGCCAGGGGTTACATTTCTTCAGAGAA 597									
QY 421 TCTGTTTCTCATCAATCTGGGATCACAAACCCAAAGCAATCCCAACCATCTACCGTGAA 480									
Db 598 TCTGTTTCTCATCAATCTGGGATCACAAACCCAAAGCAATCCCAACCATCTACCGTGAA 657									
QY 481 CTTAGGTTCTACATAGGCTTCAGAGGTGTCATTTGCAAGATGATTCGAAACTCT 540									
Db 658 CTTAGGTTCTACATAGGCTTCAGAGGTGTCATTTGCAAGATGATTCGAAACTCT 717									

```

RESULT 4
AY133811
LOCUS
DEFINITION
Arabidopsis thaliana clone U1999 putative phosphoethanolamine
N-methyltransferase (At1g48600) mRNA, complete cds.
ACCESSION
AY133811
VERSION
AY133811.1 GI:22136801
KEYWORDS
FLI CDNA.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 1459)
Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Deng,J.M., Huan,V.W.,
Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Wallender,B.K., Wong,C.,
Wu,H.C., Yu,G., Yuan,S., Bowser,L., Carninci,P., Chen,H., Cheuk,R.,
Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G.,
Kawai,J., Kim,C., Lam,B., Lin,J., Miranda,M., Narusaka,M.,
Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shim,P.,
Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
Arabidopsis Open Reading Frame (ORF) Clones
Unpublished
2 (bases 1 to 1459)
Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Deng,J.M., Huan,V.W.,
Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Wallender,B.K.,
Wong,C., Wu,H.C., Yu,G., Yuan,S., Bowser,L., Carninci,P., Chen,H.,
Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A.,
Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Miranda,M.,
Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M.,
Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and
Theologis,A.
Direct Submission
Submitted (17-JUL-2002) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
The RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.
The Salk, Stanford, PGEC (SSP) Consortium members constructed and
sequenced the PUNI (ORF) clones using the RAFL cDNAs: Yamada,K.,
Chan,M.M., Chang,C.H., Dale,J.M., Deng,J.M., Huan,V.W., Lee,J.M.,
Quach,H.L., Tang,C.C., Toriumi,M., Wallender,B.K., Wong,C.,
Wu,H.C., Yu,G., Yuan,S., Bowser,L., Carninci,P., Chen,H., Cheuk,R.,
Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G.,
Kawai,J., Kim,C., Lam,B., Lin,J., Miranda,M., Narusaka,M.,
Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shim,P.,
Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and
Theologis,A.
Yamada,K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally
to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP
/PGEC) contributed equally to this work as PIs.
Annotation is based on the January 2002 version of the Arabidopsis
genome submitted to GenBank.
FEATURES
source
1..1459
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="U11999"
/ecotype="Columbia"
note="This clone is in PUNI 51."
1..1459
/gene="At1g48600"
1..1428
/gene="At1g48600"
/codon_start=1
/evidence=experimental
/product="putative phosphoethanolamine"
gene
CDS

```

```

N-methyltransferase"
/protein_id="AA091745.1"
/db_xref="GI:22136802"
/translation="MEHSSDUTVAMMLDKASDLKEKPEVLSLIPYRGKSVLEL
GAGIKRFTGLAQAGEVIALDFTESAIQKNESVNGHYKNIKPMCADVTSPLDKIKDG
SIDLIPSNMLLYLSDKEVLMALMIGVYKPGGYTFFRSCFHQSGDKRSKNPHTY
REPRFTYKVFQECOTDRASGNSPGLSVGCKICGAYVKNQKQNCQICWIKQKVSVD
KDFRFLDNVOYKSGSGLLYRVERFVGSTGGFTTKFPAKMDLKPQKQVLDVDCG
ICGGDYMAENFDFHVVICIDLSVNMISFALERAIGLKCSVEFEVADCTTKTYDNGSD
VIYSRDTLTIHQDKPALPFTFFKVLKPGKVLITDYCRSAETSPSEAEYIKRGYDL
HGVQYCGMLNDAGDFDDVIAEDRTDQVQLRELEKVEKEKEEPISEDSSESDYDII
GWSAKLRTASGSGKGLFIADKK"
misc_difference 282
/gene="At1g48600"
/note="compared to genomic sequence"
/replace="t"
1429..1459
/gene="At1g48600"

3'UTR
ORIGIN
Query Match 2.2% Score 35; DB 8; Length 1459;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 44 AGCTTCTGATTTGACAAAGAACGCTCTGAG 78
Db 50 AAGCTTCTGATTTGACAAAGAACGCTCTGAG 84

RESULT 5
AF428454
LOCUS
DEFINITION
Arabidopsis thaliana At1g48600/TIN15_20 mRNA, complete cds.
ACCESSION
AF428454
VERSION
AF428454.1 GI:16226648
KEYWORDS
FLI CDNA.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 1656)
Chouk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,
Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,
Davis,R.W., Theologis,A. and Ecker,J.R.
Arabidopsis cDNA clones
Unpublished
2 (bases 1 to 1656)
Chouk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,
Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,
Davis,R.W., Theologis,A. and Ecker,J.R.
Direct Submission
Submitted (05-OCT-2001) Salk Institute Genomic Analysis Laboratory
(SIGNAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.
The Salk, Stanford, PGEC (SSP) Consortium members carried out the

```

sequencing and annotation of the RAFL cDNAs: Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Shinn, P., Banh, J., Bowser, L., Dale, J.M., Goldsmith, A.D., Jiang, P.X., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W., Theologis, A., and Ecker, J.R.

Cheuk, R. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

Location/Qualifiers  
1. .1656  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="1"  
/clone="RAFL09-71-K18(R14923)"  
/ecotype="Columbia"  
1. .1446  
/notes="putative phosphoethanolamine N-methyltransferase"  
/codon\_start=1  
/product="At1g48600/TIN15\_20"  
/protein\_id="AAL16223.1"  
/db\_xref="GI:16226649"  
/translation="MEHSSDLTVEAMLDKASDLKKEERPEVLISLIPPYEGKSVLEL  
GAGIGRFTGELAQAGEVIALDTESALQKNSVNGHYKNIKPCADVTSPDLKIKDG  
SIDLIFSNWLLMYLSDKEVLAEMRMIGWPGGYIFFRESCHQSGDSKRKNPHTY  
KDFQRFIDNVQKSGILRYERFVGGVSTGFTTKEFVAKMDLPGQKVLVDGCG  
IGGDFPMAENFDVHVVDLISVNI SPALERAIGLKCSVEFVADCTTKTYPDSFD  
VIYSRDTLHIQDKALPRTFFKMLKPGKVLITDYCRSAETSPSPAEYIKQGYDL  
HDVQAYGQMLKAGFDVIAEDRTDQFVQLRLELKEKEKEBEFISDFSEEDYNDIV  
GWSAKLERTASGEQKWLFIADKK"  
1447. .1656

5'UTR  
CDS

ORIGIN

Query Match 2.2%; Score 35; DB 8; Length 1656;  
Best Local Similarity 100.0%; Pred. No. 3.7e-07;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 AAGCTTCTGATCTTGACAAAGAGACGCTCTGAG 78  
|||||  
DB 68 AAGCTTCTGATCTTGACAAAGAGACGCTCTGAG 102  
|||||

RESULT 6  
AY063866  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

Arabidopsis thaliana 1660 bp mRNA linear PLN 18-SEP-2002  
N-methyltransferase (At1g48600) mRNA, complete cds.  
AY063866  
AY063866.1 GI:17380783  
FLI CDNA.

Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi

1 (bases 1 to 1660)  
Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K.,  
Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L.,  
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,  
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,  
Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,  
Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K.,  
Davis, R.W., Ecker, J.R. and Theologis, A.

Arabidopsis Full Length cDNA Clones  
Unpublished.  
REFERENCE  
2 (bases 1 to 1660)  
Yamada, K., Banh, J., Banno, F., Chang, E., Dale, J.M., Goldsmith, A.D.,  
Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C.,

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Query Match 2.2%; Score 35; DB 8; Length 1660;  
Best Local Similarity 100.0%; Pred. No. 3.7e-07;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 AAGCTTCTGATCTTGACAAAGAGACGCTCTGAG 78  
|||||  
DB 68 AAGCTTCTGATCTTGACAAAGAGACGCTCTGAG 102  
|||||

RESULT 6  
AY063866  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

Arabidopsis thaliana 1660 bp mRNA linear PLN 18-SEP-2002  
N-methyltransferase (At1g48600) mRNA, complete cds.  
AY063866  
AY063866.1 GI:17380783  
FLI CDNA.

Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi

1 (bases 1 to 1660)  
Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K.,  
Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L.,  
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,  
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,  
Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,  
Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K.,  
Davis, R.W., Ecker, J.R. and Theologis, A.

Arabidopsis Full Length cDNA Clones  
Unpublished.  
REFERENCE  
2 (bases 1 to 1660)  
Yamada, K., Banh, J., Banno, F., Chang, E., Dale, J.M., Goldsmith, A.D.,  
Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C.,

Yamamura, Y., Yu, G., Yu, S., Bowser, L., Carninci, P., Chen, H.,  
Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A.,  
Karlin-Neumann, G., Kawai, J., Kim, C., Koesema, E., Lam, B., Lin, J.,  
Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,  
Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A.,  
Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

Direct Submission  
Submitted (19-NOV-2001) Plant Gene Expression Center, 800 Buchanan  
Street, Albany, CA 94710, USA

RIKEN Genomic Sciences Center (GSC) members carried out the  
collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN  
Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J.,  
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,  
Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the  
sequencing and annotation of the RAFL cDNAs: Yamada, K., Banh, J.,  
Banno, F., Chang, E., Dale, J.M., Goldsmith, A.D., Lee, J.M.,  
Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C.,  
Yamamura, Y., Yu, G., Yu, S., Bowser, L., Chen, H., Cheuk, R., Jones, T.,  
Karlin-Neumann, G., Kim, C., Koesema, E., Lam, B., Lin, J., Meyers, M.C.,  
Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A.,  
Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to  
this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC)  
contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis  
genome submitted to GenBank.

Location/Qualifiers  
1. .1660  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="1"  
/clone="RAFL06-86-D12 (R11999)"  
/ecotype="Columbia"  
/note="This clone is in a modified pBluescript vector  
(Lambda ZAP) as a XhoI/SstI insert."  
1. .1660  
/genes="At1g48600"  
1. .34  
/genes="At1g48600"  
35. .1462  
/genes="At1g48600"  
/codon\_start=1  
/evidence="experimental  
N-methyltransferase"  
/product="putative phosphoethanolamine  
N-methyltransferase"  
/protein\_id="AAL16222.1"  
/db\_xref="GI:17380784"  
/translation="MEHSSDLTVEAMLDKASDLKKEERPEVLISLIPPYEGKSVLEL  
GAGIGRFTGELAQAGEVIALDTESALQKNSVNGHYKNIKPCADVTSPDLKIKDG  
SIDLIFSNWLLMYLSDKEVLAEMRMIGWPGGYIFFRESCHQSGDSKRKNPHTY  
KDFQRFIDNVQKSGILRYERFVGGVSTGFTTKEFVAKMDLPGQKVLVDGCG  
IGGDFPMAENFDVHVVDLISVNI SPALERAIGLKCSVEFVADCTTKTYPDSFD  
VIYSRDTLHIQDKALPRTFFKMLKPGKVLITDYCRSAETSPSPAEYIKQGYDL  
HDVQAYGQMLKAGFDVIAEDRTDQFVQLRLELKEKEKEBEFISDFSEEDYNDIV  
GWSAKLERTASGEQKWLFIADKK"

gene  
5'UTR  
CDS

misc\_difference 316  
/genes="At1g48600"  
/notes="compared to genomic sequence"  
/replaces="t"  
1463. .1660  
/genes="At1g48600"

3'UTR  
ORIGIN

Query Match 2.2%; Score 35; DB 8; Length 1660;  
Best Local Similarity 100.0%; Pred. No. 3.7e-07;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 AAGCTTCTGATCTTGACAAAGAGACGCTCTGAG 78  
|||||  
DB 68 AAGCTTCTGATCTTGACAAAGAGACGCTCTGAG 102  
|||||

RESULT 6  
AY063866  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

Arabidopsis thaliana 1660 bp mRNA linear PLN 18-SEP-2002  
N-methyltransferase (At1g48600) mRNA, complete cds.  
AY063866  
AY063866.1 GI:17380783  
FLI CDNA.

Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi

1 (bases 1 to 1660)  
Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K.,  
Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L.,  
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,  
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,  
Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,  
Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K.,  
Davis, R.W., Ecker, J.R. and Theologis, A.

Arabidopsis Full Length cDNA Clones  
Unpublished.  
REFERENCE  
2 (bases 1 to 1660)  
Yamada, K., Banh, J., Banno, F., Chang, E., Dale, J.M., Goldsmith, A.D.,  
Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C.,

TITLE  
JOURNAL  
REFERENCE  
AUTHORS





CDS  
F\*  
join(21967..22130,22278..22378)  
/note="unknown protein; similar to ESTs gb|AF162845.1,  
gb|AI992553.1, and gb|AA042117.1"  
/codon start=1  
/evidence=not experimental  
/product="TIN15.5"  
/protein\_id="AAF79693.1"  
/db\_xref="GI:8778685"  
/translation="MALQWLILSYVAVEVITLVLPMLKRVQLVSLIQLP  
ASIVAFQQLDLYWKAHRLSCSSCTATERDREKSVSTIP"  
complement(join(23638..2481,24564..24859,25106..25237))  
/note="unknown protein; similar to EST gb|AI996045.1"  
/codon start=1  
/evidence=not experimental  
/product="TIN15.6"  
/protein\_id="AAF79715.1"  
/db\_xref="GI:8778707"  
/translation="MDAVAMLPQSPILRLTSESLIESTPCFGRMTRSSGQPELR  
QNSLKPSPQSGTICVKSRSFYKASGSDASTESIAFLQLKSVQGLQLVLS  
HPLVPAVEQQLQIDRDABEQSKASSVLGTDIVLIRIAEYKERRALEEI  
LYALVQKFDANVTLPVPSITSSADPSGRVDTPMTLDGELRLHSPVTEMIONHLS  
IILKRLDITAVQISLGLGVQVAAVSMYGYFLKRDQFQLEKTRMLPQSGDEG  
MSEIQAGRDVERNFYEEABETTYQAVSNQDVGFVGINASGGSFDMKQSKLTYV  
BSPGSETLQRYATIRSESGVGIIEKHTALFCEPEIVITPQGTIDSSKDHKISFKG  
LKLVLAVTPQSFWDVSHVSRYPVLN"  
join(26361..26450,26546..26588,26676..27014,27366..27706,  
27823..28041)  
/note="unknown protein; similar to ESTs emb|Z33851.1,  
gb|Z33009.1, and gb|AA94999.1"  
/codon start=1  
/evidence=not experimental  
/product="TIN15.7"  
/protein\_id="AAF79694.1"  
/db\_xref="GI:8778686"  
/translation="MESKALICLFLPRLRSSPRLLSLPSPASSTFATRHKLDSR  
QTLWKPQSRVRSQSSQSDRPEKQKQSNYASAEIPLKSGSVGFNGLTQOL  
VERSLKVPSPQSGTICVKSRSFYKASGSDASTESIAFLQLKSVQGLQLVLS  
HPLVPAVEQQLQIDRDABEQSKASSVLGTDIVLIRIAEYKERRALEEI  
LYALVQKFDANVTLPVPSITSSADPSGRVDTPMTLDGELRLHSPVTEMIONHLS  
IILKRLDITAVQISLGLGVQVAAVSMYGYFLKRDQFQLEKTRMLPQSGDEG  
MSEIQAGRDVERNFYEEABETTYQAVSNQDVGFVGINASGGSFDMKQSKLTYV  
BSPGSETLQRYATIRSESGVGIIEKHTALFCEPEIVITPQGTIDSSKDHKISFKG  
LKLVLAVTPQSFWDVSHVSRYPVLN"  
join(28676..28749,28903..28942,29019..29111,29232..29273,  
29360..29466,29687..29774,29870..30073,30240..30293,  
30397..30594,30675..30818)  
/note="similar to glutamine synthetase cytosolic isozyme 2  
sp|P5119|GLN2 VITVI; similar to EST emb|Z29893.1"  
/codon start=1  
/evidence=not experimental  
/product="TIN15.8"  
/protein\_id="AAF79695.1"  
/db\_xref="GI:8778687"  
/translation="MTSPSLDLNLDSITKIIARIYIWIGSGMDIRSKARTLPGPV  
SNPTKLPKNDGSDSDQAGDDSEAFKDPKKNINILVCMCDYRAGDDIPFNH  
KAVIFDHPNVAABEFHIEQYTLKEDVWPLGFPQPGQYCAVAGADKA  
FGRIDVHAKYACLSGLSIGGAVGQWFGQFQISPTVIGIGAGDQWLVARYILERI  
AIVGSENRRLTGKHETADINTPSCGVADRGASVVRGDRTEKSGYFEDRPPSNMOP  
YLVTSMIAETIL"  
join(33367..34735,35037..35635)  
/note="similar to receptor-like protein kinase  
2.23; Score 35; DB 8; Length 95295;  
Best Local Similarity 100.0%; Pred. No. 2.7e-07;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 44 AAGCTTCTGATCTTGACAAAGAGACGCTCTGAG 78  
DB 81389 AAGCTTCTGATCTTGACAAAGAGACGCTCTGAG 81423

LOCUS AX250553 1120 bp DNA linear PAT 05-OCT-2001  
DEFINITION Sequence 3 from Patent WO0168870.  
ACCESSION AX250553  
VERSION AX250553.1 GI:15984290  
KEYWORDS  
SOURCE Spinacia oleracea (spinach)  
ORGANISM Spinacia oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Caryophyllales; Amaranthaceae; Spinacia.  
1  
REFERENCE  
AUTHORS Hanson, A.D.; Nuccio, M.L. and Henry, S.A.  
TITLE S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase  
compositions and methods for modulating lipid biosynthesis in  
plants  
JOURNAL Patent: WO 0168870-A 3 20-SEP-2001;  
UNIVERSITY OF FLORIDA (US); Carnegie-Mellon University (US)  
FEATURES  
source  
1..1120  
/organism="Spinacia oleracea"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3562"  
ORIGIN  
Query Match 2.0%; Score 32; DB 6; Length 1120;  
Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 130 GCTGTATTGCTGCTTTTACTGCTGAATTGCG 161  
DB 446 GCTGTATTGCTGCTTTTACTGCTGAATTGCG 477  
RESULT 9  
LOCUS AX250551 2235 bp DNA linear PAT 05-OCT-2001  
DEFINITION Sequence 1 from Patent WO0168870.  
ACCESSION AX250551  
VERSION AX250551.1 GI:15984289  
KEYWORDS  
SOURCE Spinacia oleracea (spinach)  
ORGANISM Spinacia oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Caryophyllales; Amaranthaceae; Spinacia.  
1  
REFERENCE  
AUTHORS Hanson, A.D.; Nuccio, M.L. and Henry, S.A.  
TITLE S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase  
compositions and methods for modulating lipid biosynthesis in  
plants  
JOURNAL Patent: WO 0168870-A 1 20-SEP-2001;  
UNIVERSITY OF FLORIDA (US); Carnegie-Mellon University (US)  
FEATURES  
source  
1..2235  
/organism="Spinacia oleracea"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3562"  
ORIGIN  
Query Match 2.0%; Score 32; DB 6; Length 2235;  
Best Local Similarity 100.0%; Pred. No. 1.4e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 130 GCTGTATTGCTGCTTTTACTGCTGAATTGCG 161  
DB 446 GCTGTATTGCTGCTTTTACTGCTGAATTGCG 477  
RESULT 10  
LOCUS AF237633 2252 bp mRNA linear PCN 10-MAY-2000  
DEFINITION Spinacia oleracea phosphoethanolamine N-methyltransferase (PEMT)  
mRNA, complete cds.

ACCESSION AF237633  
 VERSION AF237633.1 GI:7407188  
 KEYWORDS  
 SOURCE Spinacia oleracea (spinach)  
 ORGANISM  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales; Anaranthaceae; Spinacia.  
 REFERENCE 1 (bases 1 to 2252)  
 AUTHORS Nuccio, M.L., Ziemak, M.J., Henry, S.A., Weretilnyk, E.A. and Hanson, A.D.  
 TITLE cDNA cloning of phosphoethanolamine N-methyltransferase from spinach by complementation in *Schizosaccharomyces pombe* and characterization of the recombinant enzyme  
 JOURNAL J. Biol. Chem. 275 (19), 14095-14101 (2000)  
 MEDLINE 20261526  
 PUBMED 10799484  
 REFERENCE 2 (bases 1 to 2252)  
 AUTHORS Nuccio, M.L. and Hanson, A.D.  
 TITLE Direct Submission  
 JOURNAL Submitted (21-PEB-2000) Horticultural Sciences, University of Florida, P.O. Box 110690, Gainesville, FL 32611, USA  
 FEATURES  
 source  
 1..2252  
 /organism="Spinacia oleracea"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3562"  
 /cissue\_type="salinized leaf"  
 1..2252  
 /gene="PEAMT"  
 284..1738  
 /gene="PEAMT"  
 /EC\_number="2.1.1.103"  
 /function="catalyzes S-adenosylmethionine dependent N-methylation of phosphoethanolamine, phosphonomethylmethanolamine and phosphodimethylmethanolamine"  
 /codon\_start=1  
 /product="phosphoethanolamine N-methyltransferase"  
 /protein\_id="AF61950.1"  
 /db\_xref="GI:7407189"  
 /translation="MAAGAMCVLQREVFKKYYWHSVLTVEANMLDSQASDLKVE  
 RPEVLSMLPPYEGKSGVLELHAGIGRFTGAEKASQVIALDFIESVKKNESINGHYK  
 NYKFCADVTPSINISNSVYIIFSNLLMYLSDEVERLVMKLPKGGYIFPR  
 ESCFQSGDKRKNPNTHYRPFRTYIKFKECIMQDSSGNSYELSLGCKGAYVKS  
 KKNQKISLWQKVDSDKGFQFLDSQYKFNLSLYRVRVFGYVSTGSLGTTKE  
 FVSKLDLPKQKLVLDGCGIGGGDFYMAENVYDVVGVLDLSINMISPALESIGLKCA  
 VETPSADCTKQYDPSFDVYSRDTLHIDQKPALEFRSHKWLAPGKGLYSIDYCKS  
 ACTPSAEFAAVIRQGVLDHVKAYGKMLKAGFVEVIAENRTDQFIQVLKELDALE  
 QEKDFIDFSEEDYNDLVGNKAVLVRTTEGGEQQQGLFIKKM"  
 ORIGIN  
 Query Match 2.0%; Score 32; DB 8; Length 2252;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-05;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 130 GCTGGTATGGTCGTTTACTGGTGAATTGCG 161  
 Db 446 GCTGGTATGGTCGTTTACTGGTGAATTGCG 477  
 RESULT 11  
 LOCUS BC049504  
 DEFINITION Dantio rerio mRNA similar to estrogen receptor-binding fragment-associated gene 9 (cDNA clone MGC:56608 IMAGE:5915434), complete cds.  
 ACCESSION BC049504  
 VERSION BC049504.1 GI:29436967  
 KEYWORDS MGC.  
 SOURCE Dantio rerio (zebrafish)  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

# REFERENCE AUTHORS

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
 Cypriniformes; Cyprinidae; Danio.  
 1 (bases 1 to 1347)  
 Strausberg, K.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F.,  
 Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
 Skapleton, L., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,  
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,  
 Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,  
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
 Villalón, D.K., Murny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
 Fahey, J., Helton, E., Kettman, M., Madan, A., Young, A.C., Shvachenko, Y.,  
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shvachenko, Y.,  
 Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,  
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
 Butterfield, Y.S., Krzywinski, M.I., Skalek, U., Smallos, D.E.,  
 Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
 Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 22388257  
 12477932  
 2 (bases 1 to 1347)  
 Strausberg, R.  
 Direct Submission  
 Submitted (31-MAR-2003) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: Dr. Sumio Sugano  
 cDNA Library Preparation: Dr. Sumio Sugano  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Genome Sequence Centre,  
 BC Cancer Agency, Vancouver, BC, Canada  
 info@bcsc.bc.ca  
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,  
 Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,  
 Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo  
 Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven  
 Nese, Pawan Pandoh, Anna-Liisa Prabhui, Parvaneh Saeedi, Jacqueline  
 Schein, Duane Smallos, Michael Smith, Lorraine Spence, Jeff Stott,  
 Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,  
 George Yang, Scott Zuyderduyn, Marco Marra.

# TITLE

# JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE JOURNAL

# REMARK COMMENT

# FEATURES source

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAP Plate: 104 Row: 1 Column: 9  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: Hexamer frequency ORF  
 analysis, similarity but not identity to protein.  
 Location/Qualifiers  
 1..1347  
 /organism="Danio rerio"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:7955"  
 /clone="MGC:56608 IMAGE:5915434"  
 /tissue\_type="Whole body, adult male"  
 /clone\_lib="Sugano SJU adult male"  
 /lab\_host="DH10B"  
 /note="Vector: pME18S-FL3"  
 207..845  
 /codon\_start=1  
 /product="Similar to estrogen receptor-binding  
 fragment-associated gene 9"  
 /protein\_id="AA049504.1"  
 /db\_xref="GI:29436968"

# CDS

```
ORIGIN
Query Match
Best Local Similarity 100.0%; Score 31; DB 5; Length 1347;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1572 GATTATGCAAAAAAAAAAAAAAAAAAAAAA 1602
Db 1295 GATTATGCAAAAAAAAAAAAAAAAAAAAAA 1325

RESULT 12
BX324004/c
LOCUS
DEFINITION
Danio rerio clone CH211-284017, WORKING DRAFT SEQUENCE, 4 unordered
pieces.
ACCESSION
BX324004
VERSION
BX324004.3 GI:35209309
KEYWORDS
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
SOURCE
Danio rerio (zebrafish)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 138116)
McLaren,S.
Direct Submission
Submitted (23-SEP-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Sep 24, 2003 this sequence version replaced gi:30349842.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zc284017
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 137487 bases at least Q40
Consensus quality: 137601 bases at least Q30
Consensus quality: 137724 bases at least Q20
Insert size: 137816; sum-of-contigs
Quality coverage: 8.58x in Q20 bases; sum-of-contigs Quality
coverage: 8.29x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 4524: contig of 4524 bp in length
* 4525 4624: gap of 100 bp
* 4625 39932: contig of 35308 bp in length
* 39933 40032: gap of 100 bp
* 40033 101266: contig of 61234 bp in length
* 101267 101366: gap of 100 bp
* 101367 138116: contig of 36750 bp in length.
Location/Qualifiers
1..138116
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"

FEATURES
source
```

```
ORIGIN
Query Match
Best Local Similarity 100.0%; Score 31; DB 2; Length 138116;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1572 GATTATGCAAAAAAAAAAAAAAAAAAAAAA 1602
Db 62624 GATTATGCAAAAAAAAAAAAAAAAAAAAAA 62594

RESULT 13
LOCUS
DEFINITION
Danio rerio clone CH211-276D14, WORKING DRAFT SEQUENCE, 2 unordered
pieces.
ACCESSION
BX293550
VERSION
BX293550.5 GI:30089274
KEYWORDS
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
SOURCE
Danio rerio (zebrafish)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 168179)
McLaren,S.
Direct Submission
Submitted (28-MAY-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Apr 23, 2003 this sequence version replaced gi:29500500.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zc276D14
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 167777 bases at least Q40
Consensus quality: 167817 bases at least Q30
Consensus quality: 167839 bases at least Q20
Insert size: 168079; sum-of-contigs
Quality coverage: 9.06x in Q20 bases; sum-of-contigs Quality
coverage: 8.14x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
```

\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 152355: contig of 152355 bp in length  
\* 152356 152455: gap of 100 bp  
\* 152456 168179: contig of 15724 bp in length.

## FEATURES

source  
1..168179  
/organism="Danio rerio"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7955"  
/clone="CH211-276D14"  
/clone\_lib="CHORI-211"  
1..152355  
/notes="assembly fragment:00072  
fragment\_chain:1  
clone\_end:SP6  
vector\_side:left"  
152455..168179  
/notes="assembly\_fragment:01003  
fragment\_chain:1  
clone\_end:T7  
vector\_side:right"

## misc\_feature

fragment\_chain:1

## misc\_feature

152455..168179  
/notes="assembly\_fragment:01003  
fragment\_chain:1  
clone\_end:T7  
vector\_side:right"

## ORIGIN

Query Match 1.9%; Score 31; DB 2; Length 168179;  
Best Local Similarity 100.0%; Pred. No. 3.6e-05; Indels 0; Gaps 0;  
Matches 31; Conservative 0; Mismatches 0;

QY 1572 GATTATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1602  
|||||  
DB 45936 GATTATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 45966

## RESULT 14

BX005125/c  
LOCUS BX005125 172587 bp DNA linear VRT 16-OCT-2003  
DEFINITION Zebrafish DNA sequence from clone CH211-127116 in linkage group 10,  
complete sequence.

BX005125

ACCESSION BX005125.11 GI:37699970

VERSION HTG.

KEYWORDS

SOURCE

ORGANISM

Danio rerio (zebrafish)

Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
Cypriniformes; Cyprinidae; Danio.

1 (bases 1 to 172587)

Beasley H.

Direct Submission

Submitted (16-OCT-2003) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
zfsh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

On Oct 16, 2003 this sequence version replaced gi:34221835.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: zfsh-help@sanger.ac.uk

-----

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all

regions were either double-stranded or sequenced with an alternate

chemistry or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such

as compression-and-repeats; all regions were covered by at least

one plasmid subclone or more than one M13 subclone; and the

assembly was confirmed by restriction digest, except on the rare

occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession  
numbers given in the feature table with their source databases:  
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information  
on the WORMPEP database can be found at  
[http://www.sanger.ac.uk/projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/projects/C_elegans/wormpep) Clone-derived  
zebrafish pUC subclones occasionally display inconsistency over the  
length of mononucleotide A/T runs and conserved TA repeats. Where  
this is found the longest good quality representation will be  
submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat  
discovery system (Zhifong Bao and Sean Eddy, submitted), and those  
beginning 'dir' were identified by Rick Waterman (Stephen Johnson  
lab, WashU). For further information see  
[http://www.sanger.ac.uk/projects/d\\_rerio/fishmask.shtm](http://www.sanger.ac.uk/projects/d_rerio/fishmask.shtm)

CH211-127116 is from a CHORI-211 BAC library  
VECTOR: pTARBAC2.1.

## FEATURES

## source

Location/Qualifiers  
1..172587  
/organism="Danio rerio"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7955"  
/clone="CH211-127116"  
/clone\_lib="CHORI-211"

## ORIGIN

Query Match 1.9%; Score 31; DB 5; Length 172587;  
Best Local Similarity 100.0%; Pred. No. 3.6e-05; Indels 0; Gaps 0;  
Matches 31; Conservative 0; Mismatches 0;

QY 1572 GATTATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1602  
|||||  
DB 48349 GATTATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 48319

## RESULT 15

BX530060

LOCUS

DEFINITION

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
Cypriniformes; Cyprinidae; Danio.

1 (bases 1 to 175745)

McLaren, S.

Direct Submission

Submitted (23-SEP-2003) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
zfsh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

On Sep 24, 2003 this sequence version replaced gi:31408112.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: zfsh-help@sanger.ac.uk

----- Project Information

Center project name: 2K283F16

----- Summary Statistics

Assembly program: XGAP4; version 4.5

Chemistry: Dye-terminator; 100% of reads

Consensus quality: 175264 bases at least Q40

Consensus quality: 175360 bases at least Q30

Consensus quality: 175437 bases at least Q20

Insert size: 175545; sum-of-contigs

Insert size: 18158; 1.6% error; agarose-fp

Quality coverage: 9.45x in Q20 bases; sum-of-contigs Quality

Coverage: 9.06x in Q20 bases; agarose-fp

-----

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 3 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

22689: contig of 22689 bp in length  
22690: gap of 100 bp  
22790: contig of 83479 bp in length  
106268: contig of 106268 bp in length  
106369: gap of 100 bp  
175745: contig of 69377 bp in length.

FEATURES

source

1. .175745  
/organism="Danio rerio"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7955"  
/clone\_lib="DKEX-283F16"  
/clone\_lib="DanioKey"

misc\_feature

1. .22689  
/note="assembly fragment:01399  
fragment\_chain:1"

misc\_feature

22790. .106268  
/note="assembly fragment:02444  
fragment\_chain:1"

misc\_feature

106369. .175745  
/note="assembly fragment:02005.0"

ORIGIN

Query Match 1.9%; Score 31; DB 2; Length 175745;  
Best Local Similarity 100.0%; Pred.No. 3.6e-05;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1572 GATTATGCAAAAAAAAAAAAAAAAAAAAAA 1602  
|||||

DB 110763 GATTATGCAAAAAAAAAAAAAAAAAAAAAA 110793  
|||||

Search completed: August 2, 2004, 13:13:09  
Job time : 6485 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2004, 08:17:48 ; Search time 668 Seconds  
(without alignment)

10188.044 Million cell updates/sec

Title: US-10-031-331b-39

Perfect score: 1602

Sequence: 1 cacacgctgattacacat.....aaaaaaaaaaaaaaaaaaaaa 1602

Scoring table: OLIGO\_NVC

Gapop 60.0 , Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N Geneseq\_29Jan04.\*  
1: Geneseqn1980s.\*  
2: Geneseqn1990s.\*  
3: Geneseqn2000s.\*  
4: Geneseqn2001as.\*  
5: Geneseqn2001bs.\*  
6: Geneseqn2002s.\*  
7: Geneseqn2003as.\*  
8: Geneseqn2003bs.\*  
9: Geneseqn2003cs.\*  
10: Geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1602	100.0	1602	4	AAF74206 DNA encod
2	32	2.0	1120	5	AAD16798 Spinach P
3	32	2.0	2235	5	AAD16797 Spinach P
4	29	1.8	438	4	AAK56740 Human imm
5	29	1.8	483	9	AAK56740 Human imm
6	29	1.8	507	7	ABZ20190 Arabidops
7	29	1.8	536	4	ABZ20190 Group III
8	29	1.8	1267	4	AAK64288 Human imm
9	29	1.8	1267	4	AAK64288 Human imm
10	28	1.7	307	6	ABX73479 Human nov
11	28	1.7	339	4	ABL86191 Human ova
12	28	1.7	419	6	AAK60367 Human can
13	28	1.7	450	4	ABT09151 Phase-1 R
14	28	1.7	450	4	ABT09151 Phase-1 R
15	28	1.7	450	4	ABT09151 Phase-1 R
16	28	1.7	450	4	ABT09151 Phase-1 R
17	28	1.7	450	4	ABT09151 Phase-1 R
18	28	1.7	450	4	ABT09151 Phase-1 R
19	28	1.7	450	4	ABT09151 Phase-1 R
20	28	1.7	450	4	ABT09151 Phase-1 R
21	28	1.7	1417	6	ABX04582 Arabidops
22	28	1.7	2000	6	ABZ16984 Arabidops
23	28	1.7	3183	6	ABZ16984 Arabidops

C 24 28 1-7 10221 4 AAF58417 Human oli  
C 25 28 1-7 11655 4 AAK71335 Human imm  
C 26 28 1-7 45121 8 ADA02744 Human TNF  
C 27 28 1-7 45121 9 ADB72482 Human TNF  
C 28 28 1-7 45121 9 ADB72482 Human TNF  
C 29 28 1-7 45121 9 ADB72482 Human TNF  
C 30 27 1-7 42 7 ABT23108 Library c  
C 31 27 1-7 42 7 ABT23108 Library c  
C 32 27 1-7 42 7 ABT23108 Library c  
C 33 27 1-7 182 2 AAO68945 Junction  
C 34 27 1-7 260 9 ADD32852 Human mit  
C 35 27 1-7 444 6 ABL64210 Stomach c  
C 36 27 1-7 444 6 ABL64210 Stomach c  
C 37 27 1-7 444 6 ABL64210 Stomach c  
C 38 27 1-7 491 6 ABM96383 Ovary can  
C 39 27 1-7 513 8 ABO57289 Gene #288  
C 40 27 1-7 752 5 ACH34265 Human col  
C 41 27 1-7 804 4 AAL23782 DNA encod  
C 42 27 1-7 978 4 AAL23782 DNA encod  
C 43 27 1-7 978 4 AAL23782 DNA encod  
C 44 27 1-7 978 4 AAL23782 DNA encod  
C 45 27 1-7 978 4 AAL23782 DNA encod

## ALIGNMENTS

## RESULT 1

AAF74206  
ID AAF74206 standard; DNA; 1602 BP.

AC AAF74206;

DT 06-AUG-2003 (revised)

DT 02-MAY-2001 (first entry)

DE DNA encoding environmental stress tolerant protein SEQ ID 39.

DE Environmental stress resistance; salt; heat; desert; transgenic plant;  
ds.

OS Suaeda japonica.

PN WO200106006-A1.

XX 25-JAN-2001.

XX 19-JUL-2000; 2000WO-JP004862.

XX 19-JUL-1999; 99JP-00235910.

XX 24-MAR-2000; 2000JP-0085377.

XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.

XX Yamada A, Ozeki Y, Saito T;

XX WPI; 2001-147355/15.

XX P-PSDB; AAB80627.

XX Screening method to obtain DNA encoding environmental stress resistance factor, useful for producing transgenic plants resistant to environmental stress.

XX Claim 65; Page 123-125; 167pp; Japanese.

XX Polynucleotide sequences AAF74187 - AAF74218 encode proteins AAB80608 - AAB80639, which impart environmental stress resistance. The invention relates to a method for identifying DNA encoding proteins imparting environmental stress resistance. The method comprises inserting cDNA from a library originating in a salt-resistant organism into a host cell, culturing the transformants under conditions in which the untransformed host does not grow well, and selecting for viable clones. The method is useful for obtaining environmental stress resistance



CC factors. The DNA encoding proteins conferring environmental stress resistance can be used in the production of plants resistant to environmental stress, which can be cultivated in unfavourable environments such as deserts, salt damaged ground, cold regions and the oceans. They can be used for increasing the area of land covered by green plants, and desert greening and afforestation, in order to counter the effects of the increase in atmospheric carbon dioxide concentration. PCR primers AAF74219 and AAF74220 are used in an example illustrating the method of the invention. (Updated on 06-AUG-2003 to correct OS field.)	
CC	XX
SQ Sequence 1602 BP; 454 A; 279 C; 383 G; 486 T; 0 U; 0 Other;	
Query Match	
Best Local Similarity 100.0%; Score 1602; DB 4; Length 1602;	
Matches 1602; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1
DB	1
QY	61
DB	61
QY	121
DB	121
QY	181
DB	181
QY	241
DB	241
QY	301
DB	301
QY	361
DB	361
QY	421
DB	421
QY	481
DB	481
QY	541
DB	541
QY	601
DB	601
QY	661
DB	661
QY	721
DB	721
QY	781
DB	781
QY	841
DB	841
QY	901
DB	901
QY	961
DB	961
QY	1021
DB	1021
QY	1081
DB	1081
QY	1141
DB	1141
QY	1201
DB	1201
QY	1261
DB	1261
QY	1321
DB	1321
QY	1381
DB	1381
QY	1441
DB	1441
QY	1501
DB	1501
QY	1561
DB	1561
QY	1602
DB	1602
QY	1662
DB	1662
QY	1722
DB	1722
QY	1782
DB	1782
QY	1842
DB	1842
QY	1902
DB	1902
QY	1962
DB	1962
QY	2022
DB	2022
QY	2082
DB	2082
QY	2142
DB	2142
QY	2202
DB	2202
QY	2262
DB	2262
QY	2322
DB	2322
QY	2382
DB	2382
QY	2442
DB	2442
QY	2502
DB	2502
QY	2562
DB	2562
QY	2622
DB	2622
QY	2682
DB	2682
QY	2742
DB	2742
QY	2802
DB	2802
QY	2862
DB	2862
QY	2922
DB	2922
QY	2982
DB	2982
QY	3042
DB	3042
QY	3102
DB	3102
QY	3162
DB	3162
QY	3222
DB	3222
QY	3282
DB	3282
QY	3342
DB	3342
QY	3402
DB	3402
QY	3462
DB	3462
QY	3522
DB	3522
QY	3582
DB	3582
QY	3642
DB	3642
QY	3702
DB	3702
QY	3762
DB	3762
QY	3822
DB	3822
QY	3882
DB	3882
QY	3942
DB	3942
QY	4002
DB	4002
QY	4062
DB	4062
QY	4122
DB	4122
QY	4182
DB	4182
QY	4242
DB	4242
QY	4302
DB	4302
QY	4362
DB	4362
QY	4422
DB	4422
QY	4482
DB	4482
QY	4542
DB	4542
QY	4602
DB	4602
QY	4662
DB	4662
QY	4722
DB	4722
QY	4782
DB	4782
QY	4842
DB	4842
QY	4902
DB	4902
QY	4962
DB	4962
QY	5022
DB	5022
QY	5082
DB	5082
QY	5142
DB	5142
QY	5202
DB	5202
QY	5262
DB	5262
QY	5322
DB	5322
QY	5382
DB	5382
QY	5442
DB	5442
QY	5502
DB	5502
QY	5562
DB	5562
QY	5622
DB	5622
QY	5682
DB	5682
QY	5742
DB	5742
QY	5802
DB	5802
QY	5862
DB	5862
QY	5922
DB	5922
QY	5982
DB	5982
QY	6042
DB	6042
QY	6102
DB	6102
QY	6162
DB	6162
QY	6222
DB	6222
QY	6282
DB	6282
QY	6342
DB	6342
QY	6402
DB	6402
QY	6462
DB	6462
QY	6522
DB	6522
QY	6582
DB	6582
QY	6642
DB	6642
QY	6702
DB	6702
QY	6762
DB	6762
QY	6822
DB	6822
QY	6882
DB	6882
QY	6942
DB	6942
QY	7002
DB	7002
QY	7062
DB	7062
QY	7122
DB	7122
QY	7182
DB	7182
QY	7242
DB	7242
QY	7302
DB	7302
QY	7362
DB	7362
QY	7422
DB	7422
QY	7482
DB	7482
QY	7542
DB	7542
QY	7602
DB	7602
QY	7662
DB	7662
QY	7722
DB	7722
QY	7782
DB	7782
QY	7842
DB	7842
QY	7902
DB	7902
QY	7962
DB	7962
QY	8022
DB	8022
QY	8082
DB	8082
QY	8142
DB	8142
QY	8202
DB	8202
QY	8262
DB	8262
QY	8322
DB	8322
QY	8382
DB	8382
QY	8442
DB	8442
QY	8502
DB	8502
QY	8562
DB	8562
QY	8622
DB	8622
QY	8682
DB	8682
QY	8742
DB	8742
QY	8802
DB	8802
QY	8862
DB	8862
QY	8922
DB	8922
QY	8982
DB	8982
QY	9042
DB	9042
QY	9102
DB	9102
QY	9162
DB	9162
QY	9222
DB	9222
QY	9282
DB	9282
QY	9342
DB	9342
QY	9402
DB	9402
QY	9462
DB	9462
QY	9522
DB	9522
QY	9582
DB	9582
QY	9642
DB	9642
QY	9702
DB	9702
QY	9762
DB	9762
QY	9822
DB	9822
QY	9882
DB	9882
QY	9942
DB	9942
QY	10002
DB	10002
QY	10062
DB	10062
QY	10122
DB	10122
QY	10182
DB	10182
QY	10242
DB	10242
QY	10302
DB	10302
QY	10362
DB	10362
QY	10422
DB	10422
QY	10482
DB	10482
QY	10542
DB	10542
QY	10602
DB	10602
QY	10662
DB	10662
QY	10722
DB	10722
QY	10782
DB	10782
QY	10842
DB	10842
QY	10902
DB	10902
QY	10962
DB	10962
QY	11022
DB	11022
QY	11082
DB	11082
QY	11142
DB	11142
QY	11202
DB	11202
QY	11262
DB	11262
QY	11322
DB	11322
QY	11382
DB	11382
QY	11442
DB	11442
QY	11502
DB	11502
QY	11562
DB	11562
QY	11622
DB	11622
QY	11682
DB	11682
QY	11742
DB	11742
QY	11802
DB	11802
QY	11862
DB	11862
QY	11922
DB	11922
QY	11982
DB	11982
QY	12042
DB	12042
QY	12102
DB	12102
QY	12162
DB	12162
QY	12222
DB	12222
QY	12282
DB	12282
QY	12342
DB	12342
QY	12402
DB	12402
QY	12462
DB	12462
QY	12522
DB	12522
QY	12582
DB	12582
QY	12642
DB	12642
QY	12702
DB	12702
QY	12762
DB	12762
QY	12822
DB	12822
QY	12882
DB	12882
QY	12942
DB	12942
QY	13002
DB	13002
QY	13062
DB	13062
QY	13122
DB	13122
QY	13182
DB	13182
QY	13242
DB	13242
QY	13302
DB	13302
QY	13362
DB	13362
QY	13422
DB	13422
QY	13482
DB	13482
QY	13542
DB	13542
QY	13602
DB	13602
QY	13662
DB	13662
QY	13722
DB	13722
QY	13782
DB	13782
QY	13842
DB	13842
QY	13902
DB	13902
QY	13962
DB	13962
QY	14022
DB	14022
QY	14082
DB	14082
QY	14142
DB	14142
QY	14202
DB	14202
QY	14262
DB	14262
QY	14322
DB	14322
QY	14382
DB	14382
QY	14442
DB	14442
QY	14502
DB	14502
QY	14562
DB	14562
QY	14622
DB	14622
QY	14682
DB	14682
QY	14742
DB	14742
QY	14802
DB	14802
QY	14862
DB	14862
QY	14922
DB	14922
QY	14982
DB	14982
QY	15042
DB	15042
QY	15102
DB	15102
QY	15162
DB	15162
QY	15222
DB	15222
QY	15282
DB	15282
QY	15342
DB	15342
QY	15402
DB	15402
QY	15462
DB	15462
QY	15522
DB	15522
QY	15582
DB	15582
QY	15642
DB	15642
QY	15702
DB	15702
QY	15762
DB	15762
QY	15822
DB	15822
QY	15882
DB	15882
QY	15942
DB	15942
QY	16002
DB	16002
QY	16062
DB	16062
QY	16122
DB	16122
QY	16182
DB	16182
QY	16242
DB	16242
QY	16302
DB	16302
QY	16362
DB	16362
QY	16422
DB	16422
QY	16482
DB	16482
QY	16542
DB	16542
QY	16602
DB	16602
QY	16662
DB	16662
QY	16722
DB	16722
QY	16782
DB	16782
QY	16842
DB	16842
QY	16902
DB	16902
QY	16962
DB	16962
QY	17022
DB	17022
QY	17082
DB	17082
QY	17142
DB	17142
QY	17202
DB	17202
QY	17262
DB	17262
QY	17322
DB	17322
QY	17382
DB	17382
QY	17442
DB	17442
QY	17502
DB	17502
QY	17562
DB	17562
QY	17622
DB	17622
QY	17682
DB	17682
QY	17742
DB	17742
QY	17802
DB	17802
QY	17862
DB	17862
QY</	

FT /\*tag= a  
 FT /product= "Spinach PEAMT truncated protein"  
 FT /note= "CDS does not include stop codon"  
 XX /partial  
 XX

PN WO200168870-A2.

XX 20-SEP-2001.

XX 15-MAR-2001; 2001WO-US008352.

XX 15-MAR-2000; 2000US-00525885.

XX (UYFL ) UNIV FLORIDA.

XX (UYCA-) UNIV CARNEGIE MELLON.

XX Hanson AD, Nuccio ML, Henry SA;

XX WPI; 2001-565796/63.

XX P-PSDB; AAE09761.

XX New plant S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase  
 XX polypeptides, useful for modulating the levels of cellular intermediates  
 XX such as phosphodimethylethanolamine and for altering the lipid content in  
 XX plants cells.

XX Claim 18; Page 109; 158pp; English.

XX The present sequence is spinach S-adenosyl-L-methionine:phospho-  
 XX ethanolamine N-methyltransferase (PEAMT) truncated DNA. The PEAMT  
 XX sequences are useful for modulating the levels of cellular intermediates  
 XX such as phosphodimethylethanolamine, phosphomono-methylethanolamine,  
 XX choline, phosphocholine, phosphatidylcholine, choline-O-sulphate or  
 XX glycine betaine. They are useful for altering the lipid content in plant  
 XX cells. The polynucleotides are also useful for improving the osmotic  
 XX stress tolerance of a plant and increasing the cryoprotectant properties  
 XX of a plant. The present invention also relates to methods and  
 XX compositions comprising PEAMT used for generating transgenic plants with  
 XX increased nutritional value

XX Sequence 1120 BP; 304 A; 218 C; 256 G; 342 T; 0 U; 0 Other;

XX Query Match 2.0%; Score 32; DB 5; Length 1120;

XX Best Local Similarity 100.0%; Pred. No. 0.0058;

XX Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GCTGGTATTGGTCGTTTACTGGTGAATTGGC 161

Db 446 GCTGGTATTGGTCGTTTACTGGTGAATTGGC 477

RESULT 3

AD16797

ID AAD16797 standard; DNA; 2235 BP.

XX AC

XX AD16797;

XX 29-NOV-2001 (first entry)

XX Spinach PEAMT DNA.

XX Spinach; S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase;  
 XX PEAMT; cellular intermediate; phospho-dimethylethanolamine; choline;  
 XX phosphomono-methylethanolamine; phosphocholine; phosphatidylcholine;  
 XX glycine betaine; choline-O-sulphate; lipid content alteration;  
 XX osmotic stress tolerance; nutritional value; transgenic plant;  
 XX cryoprotectant; ds.

XX Spinacia oleracea.

XX Key

XX Location/Qualifiers

FT 254..1738

FT CDS /\*tag= a

FT /product= "Spinach PEAMT protein"

XX WO200168870-A2.

XX 20-SEP-2001.

XX 15-MAR-2001; 2001WO-US008352.

XX 15-MAR-2000; 2000US-00525885.

XX (UYFL ) UNIV FLORIDA.

XX (UYCA-) UNIV CARNEGIE MELLON.

XX Hanson AD, Nuccio ML, Henry SA;

XX WPI; 2001-565796/63.

XX P-PSDB; AAE09760.

XX New plant S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase  
 XX polypeptides, useful for modulating the levels of cellular intermediates  
 XX such as phosphodimethylethanolamine and for altering the lipid content in  
 XX plants cells.

XX Claim 18; Page 108-109; 158pp; English.

XX The present sequence is a DNA encoding spinach S-adenosyl-L-methionine:  
 XX phosphoethanolamine N-methyltransferase (PEAMT). The PEAMT sequences are  
 XX useful for modulating the levels of cellular intermediates such as  
 XX phosphodimethylethanolamine, phosphomono-methylethanolamine, choline,  
 XX phosphocholine, phosphatidylcholine, choline-O-sulphate or glycine  
 XX betaine. They are useful for altering the lipid content in plant cells.  
 XX The polynucleotides are also useful for improving the osmotic stress  
 XX tolerance of a plant and increasing the cryoprotectant properties of a  
 XX plant. The present invention also relates to methods and compositions  
 XX comprising PEAMT used for generating transgenic plants with increased  
 XX nutritional value

XX Sequence 2235 BP; 623 A; 399 C; 516 G; 697 T; 0 U; 0 Other;

XX Query Match 2.0%; Score 32; DB 5; Length 2235;

XX Best Local Similarity 100.0%; Pred. No. 0.0052;

XX Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GCTGGTATTGGTCGTTTACTGGTGAATTGGC 161

Db 446 GCTGGTATTGGTCGTTTACTGGTGAATTGGC 477

RESULT 4

AAK56740

ID AAK56740 standard; cDNA; 438 BP.

XX AC

XX AAK56740;

XX 06-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:1800.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

XX cytostatic; gene therapy; vaccine; metastasis; ss.

XX Homo sapiens.

XX WO200157182-A2.

XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US001354.

XX 31-JAN-2000; 2000US-0179085P.

XX 04-FEB-2000; 2000US-0180628P.

XX 24-FEB-2000; 2000US-0184664P.

XX 02-MAR-2000; 2000US-0186350P.

PR	16-MAR-2000;	2000US-0189874P.	PR	13-OCT-2000;	2000US-0239935P.
PR	17-MAR-2000;	2000US-0190076P.	PR	13-OCT-2000;	2000US-0239937P.
PR	18-APR-2000;	2000US-0198123P.	PR	20-OCT-2000;	2000US-0240360P.
PR	19-MAY-2000;	2000US-0205515P.	PR	20-OCT-2000;	2000US-0241221P.
PR	07-JUN-2000;	2000US-0209467P.	PR	20-OCT-2000;	2000US-0241785P.
PR	28-JUN-2000;	2000US-0214886P.	PR	20-OCT-2000;	2000US-0241786P.
PR	30-JUN-2000;	2000US-0215135P.	PR	20-OCT-2000;	2000US-0241787P.
PR	07-JUL-2000;	2000US-0216647P.	PR	20-OCT-2000;	2000US-0241808P.
PR	11-JUL-2000;	2000US-0217487P.	PR	20-OCT-2000;	2000US-0241809P.
PR	11-JUL-2000;	2000US-0217496P.	PR	01-NOV-2000;	2000US-0241826P.
PR	14-JUL-2000;	2000US-0218290P.	PR	08-NOV-2000;	2000US-024617P.
PR	26-JUL-2000;	2000US-0220963P.	PR	08-NOV-2000;	2000US-0246474P.
PR	26-JUL-2000;	2000US-0220964P.	PR	08-NOV-2000;	2000US-0246475P.
PR	14-AUG-2000;	2000US-0224518P.	PR	08-NOV-2000;	2000US-0246476P.
PR	14-AUG-2000;	2000US-0224519P.	PR	08-NOV-2000;	2000US-0246477P.
PR	14-AUG-2000;	2000US-0225213P.	PR	08-NOV-2000;	2000US-0246523P.
PR	14-AUG-2000;	2000US-0225214P.	PR	08-NOV-2000;	2000US-0246524P.
PR	14-AUG-2000;	2000US-0225256P.	PR	08-NOV-2000;	2000US-0246525P.
PR	14-AUG-2000;	2000US-0225257P.	PR	08-NOV-2000;	2000US-0246526P.
PR	14-AUG-2000;	2000US-0225258P.	PR	08-NOV-2000;	2000US-0246528P.
PR	14-AUG-2000;	2000US-0225270P.	PR	08-NOV-2000;	2000US-0246532P.
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246609P.
PR	14-AUG-2000;	2000US-0225757P.	PR	08-NOV-2000;	2000US-0246610P.
PR	14-AUG-2000;	2000US-0225758P.	PR	08-NOV-2000;	2000US-0246611P.
PR	14-AUG-2000;	2000US-0225759P.	PR	08-NOV-2000;	2000US-0246613P.
PR	18-AUG-2000;	2000US-0226279P.	PR	17-NOV-2000;	2000US-0249207P.
PR	22-AUG-2000;	2000US-0226681P.	PR	17-NOV-2000;	2000US-0249208P.
PR	22-AUG-2000;	2000US-0226868P.	PR	17-NOV-2000;	2000US-0249209P.
PR	22-AUG-2000;	2000US-0227182P.	PR	17-NOV-2000;	2000US-0249210P.
PR	23-AUG-2000;	2000US-0227182P.	PR	17-NOV-2000;	2000US-0249211P.
PR	10-AUG-2000;	2000US-0228924P.	PR	17-NOV-2000;	2000US-0249212P.
PR	01-SEP-2000;	2000US-0229287P.	PR	17-NOV-2000;	2000US-0249213P.
PR	01-SEP-2000;	2000US-0229343P.	PR	17-NOV-2000;	2000US-0249214P.
PR	01-SEP-2000;	2000US-0229344P.	PR	17-NOV-2000;	2000US-0249215P.
PR	01-SEP-2000;	2000US-0229345P.	PR	17-NOV-2000;	2000US-0249216P.
PR	05-SEP-2000;	2000US-0229509P.	PR	17-NOV-2000;	2000US-0249217P.
PR	05-SEP-2000;	2000US-0229513P.	PR	17-NOV-2000;	2000US-0249218P.
PR	06-SEP-2000;	2000US-0230437P.	PR	17-NOV-2000;	2000US-0249244P.
PR	06-SEP-2000;	2000US-0230438P.	PR	17-NOV-2000;	2000US-0249245P.
PR	08-SEP-2000;	2000US-0231242P.	PR	17-NOV-2000;	2000US-0249264P.
PR	08-SEP-2000;	2000US-0231243P.	PR	17-NOV-2000;	2000US-0249265P.
PR	08-SEP-2000;	2000US-0231244P.	PR	17-NOV-2000;	2000US-0249297P.
PR	08-SEP-2000;	2000US-0231413P.	PR	17-NOV-2000;	2000US-0249299P.
PR	08-SEP-2000;	2000US-0231414P.	PR	17-NOV-2000;	2000US-0249300P.
PR	08-SEP-2000;	2000US-0232080P.	PR	01-DEC-2000;	2000US-0250160P.
PR	08-SEP-2000;	2000US-0232081P.	PR	01-DEC-2000;	2000US-0250391P.
PR	12-SEP-2000;	2000US-0231968P.	PR	05-DEC-2000;	2000US-0251030P.
PR	14				

CC amino acid sequences given in AAK821170 to AAK91921. (I) have cytostatic  
 CC activity, and can be used in gene therapy and vaccine production. (I)  
 CC proteins and polynucleotides may be used in the prevention, diagnosis and  
 CC treatment of diseases associated with inappropriate (I) expression. For  
 CC example, they may be used to treat disorders associated with decreased  
 CC expression by rectifying mutations or deletions in a patient's genome  
 CC that affect the activity of (I) by expressing inactive proteins or to  
 CC supplement the patient's own production of (I). Additionally, (I)  
 CC polynucleotides may be used to produce the secreted (I), by inserting the  
 CC nucleic acids into a host cell and culturing the cell to express the  
 CC protein. (I) proteins and polynucleotides may be used to prevent,  
 CC diagnose and treat immune/haematopoietic-related diseases, especially  
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
 CC to AAK87694 represent human immune/haematopoietic antigen genomic  
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK92169  
 CC represent sequences used in the exemplification of the present invention  
 XX

SQ Sequence 438 BP; 177 A; 63 C; 67 G; 128 T; 0 U; 3 Other;  
 Query Match 1.84; Score 29; DB 4; Length 438;  
 Best Local Similarity 100.0%; Pred. No. 0.1;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1574 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 1602  
 Db 397 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 425

## RESULT 5

ID ADE82102 standard; cDNA; 483 BP.

XX AC ADE82102;

XX DT 29-JAN-2004 (first entry)

XX DE Arabidopsis thaliana expressed polynucleotide seq id 873.

XX KW Genetically modified organism; transgenic organism; plant;  
 KW inhibitor testing; activator testing; modifier testing; fungicide;  
 KW insecticide; genetic function; genetic regulation; cellular metabolism;  
 KW gene; ss.

XX OS Arabidopsis thaliana.

XX FN US2003115639-A1.

XX PD 19-JUN-2003.

XX PF 26-JAN-2001; 2001US-00770961.

XX PR 27-JAN-2000; 2000US-0178466P.

XX PA (GORL/) GORLACH J.

XX PA (ANY/) AN Y.

XX PA (HAMI/) HAMILTON C M.

XX PA (PRIC/) PRICE J L.

XX PA (RAIN/) RAINES T M.

XX PA (YUY/) YU Y.

XX PA (RAME/) RAMEKA J G.

XX PA (PAGE/) PAGE A.

XX PA (MATH/) MATHW A V.

XX PA (LEDF/) LEDFORD B L.

XX PA (WOES/) WOESSNER J P.

XX PA (HAAS/) HAAS W D.

XX PA (GARC/) GARCIA C A.

XX PA (KRIC/) KRICKER M.

XX PA (SLAT/) SLATER T.

XX PA (DAV/) DAVIS K R.

XX PA (ALLE/) ALLEN K.

XX PA (HOFF/) HOFFMAN N.

XX PA (HURB/) HURBAN P.

PI Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y,  
 PI Rameka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD,  
 PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;  
 PI Hurban P;  
 XX WPI; 2003-810930/76.

XX Novel Arabidopsis thaliana nucleic acids useful for generating  
 PT genetically modified transgenic organisms, for screening biologically  
 PT active agents such as fungicides, insecticides.

XX Claim 1; SEQ ID NO 873; 44pp; English.

XX The invention describes a nucleic acid (I) comprising a sequence capable  
 CC of hybridising under stringent conditions to any one of 999 fully defined  
 CC Arabidopsis thaliana sequences (I) as given in specification e.g., 360,  
 CC 1137, 455, 219, 472 nucleotides, etc, or its fragment. (I) is useful as a  
 CC hybridisation probe to complementary molecules in a cDNA library. (I) is  
 CC also useful for generating genetically modified and transgenic organisms,  
 CC usually plant cells and plants. A protein encoded by (I) is useful in  
 CC screening assays to determine the effect of candidate inhibitors,  
 CC activators or modifiers of the gene product. The protein is also useful  
 CC for screening biologically active agents e.g., fungicides and  
 CC insecticides. A genetically modified cell, comprising an exogenous  
 CC nucleic acid, where the nucleic acid comprises transcription regulatory  
 CC sequences operably linked to a sequence capable of hybridising under  
 CC stringent conditions to (I) is useful in the study of genetic function  
 CC and regulation, for alteration of the cellular metabolism and for  
 CC screening compounds that may affect the biological function of the gene  
 CC or gene product. This sequence represents an Arabidopsis thaliana  
 CC polynucleotide of the invention.

SQ Sequence 483 BP; 142 A; 115 C; 76 G; 150 T; 0 U; 0 Other;

Query Match 1.84; Score 29; DB 9; Length 483;  
 Best Local Similarity 100.0%; Pred. No. 0.1;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1574 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 1602

Db 48 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 20

## RESULT 6

ABZ20190/C

ID ABZ20190 standard; cDNA; 507 BP.

XX AC ABZ20190;

XX DT 23-JAN-2003 (first entry)

XX DE Group III cDNA cancer related clone SEQ ID NO:2616.

XX KW Human; cancer; tumour; therapy; diagnosis; CT antigen; CP antigen;

XX KW immune response; virology; immunology; microbiology; molecular biology;

XX KW recombinant DNA technology; gene; ss.

XX OS Homo sapiens.

XX PN WO200278516-A2..

XX PD 10-OCT-2002.

XX PF 28-MAR-2002; 2002WO-US010421.

XX PR 30-MAR-2001; 2001US-0280255P.

XX PR 28-AUG-2001; 2001US-0315563P.

XX PR 09-JAN-2002; 2002US-0347313P.

XX PA (CORI-) CORIXA CORP.

XX PI Wang T, Wang S, Bangur CS, Gaiger A;

```

DR WPI; 2003-058387/05.
XX New immunogenic polynucleotides or polypeptides useful for diagnosing,
PT preventing and treating cancer expressing CT or Cp mRNA antigens, and in
PT virology, immunology, microbiology, molecular biology and recombinant DNA
PT techniques.
XX Claim 1; SEQ ID NO 2616; 207pp; English.
XX ABQ17575 to ABQ20506 represent isolated polynucleotide (I) sequences, and
CC ABP54446 to ABP54472 represent protein (II) sequences, from the present
CC invention. (I) and (II) have cytostatic activity and can be used in gene
CC therapy and vaccines. (I), (II), antibodies and compositions from the
CC present invention are useful for diagnosing, preventing and treating
CC cancer, which expresses CT or Cp mRNA antigens. They are useful for
CC stimulating immune response. They can also be useful in virology,
CC immunology, microbiology, molecular biology and recombinant DNA
CC techniques. N.B. The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at fcp.wipo.int/pub/published_pct_sequences
XX Sequence 507 BP; 154 A; 79 C; 87 G; 175 T; 0 U; 12 Other;
SQ Query Match 1.8%; Score 29; DB 7; Length 507;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1574 TTTATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1602
DB 34 TTTATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 6
RESULT 7
AAK64289/c
ID AAK64289 standard; cDNA; 536 BP.
XX AAK64289;
AC AAK64289;
DT 06-NOV-2001 (first entry)
DE Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:9349.
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW Cytostatic; gene therapy; vaccine; metastasis; ss.
XX Homo sapiens.
OS WO200157182-A2.
XX PD 09-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001354.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218220P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229511P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234597P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.

```

PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249276P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251039P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 06-DEC-2000; 2000US-0256719P.  
PR 08-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 11-DEC-2000; 2000US-0251990P.  
PR 05-JAN-2001; 2001US-0259678P.  
PR XX  
PR PA (HUMA-) HUMAN GENOME SCI INC.  
PR XX  
PR PI Rosen CA, Barash SC, Ruben SM;  
PR XX  
PR DR WPI; 2001-483426/52.  
PR DR P-PSDB; AAM91508.  
PR XX  
PR FT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PR FT useful for preventing, diagnosing and/or treating cancers and metastasis.  
PR XX  
PR XX Claim 1; SEQ ID NO 9349; 3071pp + Sequence Listing; English.  
PR XX  
PR CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
CC amino acid sequences given in AAM91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC treatments and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting the  
CC nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/hematopoietic-related diseases, especially  
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/hematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
CC represent sequences used in the exemplification of the present invention  
CC

XX  
SQ Sequence 536 BP; 167 A; 97 C; 89 G; 182 T; 0 U; 1 Other;  
Query Match 1.8%; Score 29; DB 4; Length 536;  
Best Local Similarity 100.0%; Pred. No. 0.1;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1574 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 1602  
Db 29 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 1  
RESULT 8  
AAS26138  
ID AAS26138 standard; cDNA; 1267 BP.  
XX  
AC AAS26138;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human cDNA encoding a novel secreted protein, Seq ID 317.  
XX  
KW Human; immunosuppressive; antiarthritic; ss; antirheumatic; cytostatic;  
KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;  
KW antibacterial; virucide; fungicide; ophthalmological; vulnery;  
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;  
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;  
KW cerebral ischaemia; angiogenesis; nervous system disorder;  
KW Alzheimer's disease; infection; ocular disorder; corneal infection;  
KW wound healing; epithelial cell proliferation; skin ageing; food additive;  
KW preservative; antiproliferative.  
XX  
OS Homo sapiens.  
XX  
PN WC20015322-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001341.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 11-JUL-2000; 2000US-0216800P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 14-AUG-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226688P.

PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231142P.  
PR 08-SEP-2000; 2000US-0231143P.  
PR 08-SEP-2000; 2000US-0231144P.  
PR 08-SEP-2000; 2000US-0231145P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0231977P.  
PR 14-SEP-2000; 2000US-0231988P.  
PR 14-SEP-2000; 2000US-0232389P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234374P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 02-OCT-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.

PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 01-DEC-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 05-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-488783/53.

P-PSDB; AAU16151.

New nucleic acid molecules encoding 461 human secreted proteins for diagnosing, preventing, treating or ameliorating medical conditions and used as food additives or preservatives.

Claim 1; SEQ ID NO 317; 980pp; English.

The invention relates to isolated nucleic acid molecules and their encoded secreted proteins. The nucleic acids and proteins are used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used in diagnosing a pathological condition or susceptibility to a pathological condition. Antibodies to the proteins can also be used in alleviating symptoms associated with the disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA). Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiodysplasia, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection. CC can also be used to aid wound healing and epithelial cell proliferation, CC to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors and other nutritional components. The present CC sequence encodes a novel secreted protein of the invention. Note: The sequence data for this patent did not form part of the printed

Query Match 1.8%; Score 29; DB 4; Length 1267;

Best Local Similarity 100.0%; Pred No 0.089;

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



QY 1574 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 1602  
 DB 1222 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 1250

## RESULT 9

ABX73479  
 ID ABX73479 standard; DNA; 1267 BP.

XX AC ABX73479;

XX DT 18-MAR-2003 (first entry)

XX DE Human novel polynucleotide #307.

XX KW Human; gene; ds; neural disorder; immune system disorder; renal disorder;  
 KW muscular disorder; respiratory disease; reproductive disorder;  
 KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;  
 KW hyperproliferative disorder; inflammatory disease; allergic reaction;  
 KW blood related disorder; cancer; immunosuppressive; antiinflammatory;  
 KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;  
 KW haemostatic; antiarteriosclerotic.

XX OS Homo sapiens.

XX FN US2002132753-A1.

XX PD 19-SEP-2002.

XX PP 17-JAN-2001; 2001US-00764864.

XX PR 31-JAN-2000; 2000US-0179065P.

XX PR 04-FEB-2000; 2000US-0180628P.

XX PR 28-JUN-2000; 2000US-0214886P.

XX PR 07-JUL-2000; 2000US-0216647P.

XX PR 07-JUL-2000; 2000US-0216880P.

XX PR 11-JUL-2000; 2000US-0217487P.

XX PR 14-JUL-2000; 2000US-0217496P.

XX PR 26-JUL-2000; 2000US-0218290P.

XX PR 26-JUL-2000; 2000US-0220963P.

XX PR 14-AUG-2000; 2000US-0220964P.

XX PR 14-AUG-2000; 2000US-0224518P.

XX PR 14-AUG-2000; 2000US-0224519P.

XX PR 14-AUG-2000; 2000US-0225267P.

XX PR 14-AUG-2000; 2000US-0225268P.

XX PR 14-AUG-2000; 2000US-0225270P.

XX PR 14-AUG-2000; 2000US-0225447P.

XX PR 14-AUG-2000; 2000US-0225757P.

XX PR 22-AUG-2000; 2000US-0225758P.

XX PR 30-AUG-2000; 2000US-0228688P.

XX PR 01-SEP-2000; 2000US-0228924P.

XX PR 01-SEP-2000; 2000US-0229287P.

XX PR 01-SEP-2000; 2000US-0229343P.

XX PR 01-SEP-2000; 2000US-0229344P.

XX PR 05-SEP-2000; 2000US-0229345P.

XX PR 05-SEP-2000; 2000US-0229509P.

XX PR 08-SEP-2000; 2000US-0229513P.

XX PR 21-SEP-2000; 2000US-0231413P.

XX PR 21-SEP-2000; 2000US-0234223P.

PR 20-OCT-2000; 2000US-0240960P.

PR 20-OCT-2000; 2000US-0241785P.

PR 01-NOV-2000; 2000US-0241809P.

PR 17-NOV-2000; 2000US-0244617P.

PR 08-DEC-2000; 2000US-0249299P.

PR 08-DEC-2000; 2000US-0251856P.

PR 08-DEC-2000; 2000US-0251868P.

XX (ROSE/) ROSEN C A.

XX (RUBE/) RUBEN S M.

XX (BARA/) BARASH S C.

XX Rosen CA, Ruben SM, Barash SC;

XX WPI, 2003-147444/14.

XX P-PSDB; ABUS5219.

XX New polypeptides and nucleic acids, useful in gene therapy for treating,

XX inhibiting or preventing e.g. neural, immune system, muscular,

XX respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or

XX renal disorders.

XX Claim 1; SEQ ID NO 317; 402pp; English.

XX The invention relates to human novel polypeptides and their associated

XX polynucleotides. The polypeptides and polynucleotides are useful in gene

XX therapy for treating, inhibiting or preventing neural disorders, immune

XX system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis

XX and multiple sclerosis), muscular disorders, respiratory diseases (e.g.

XX nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,

XX gastrointestinal disorders, pulmonary disorders, cardiovascular disorders,

XX (e.g. congenital heart defects, Epstein's anomaly and hypoplastic left

XX heart syndrome), renal disorders (e.g. acute kidney failure and end-stage

XX renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and

XX leukaemia), inflammatory diseases (e.g. septic shock, bursitis and

XX appendicitis), allergic reactions and conditions (e.g. asthma), blood

XX related disorders (e.g. thrombosis, atherosclerosis and myocardial

XX infarction) and cancerous diseases. Sequences ABX73173-ABX74157 represent

XX human novel polynucleotides of the invention

SQ Sequence 1267 BP; 316 A; 386 C; 317 G; 248 T; 0 U; 0 Other;

Query Match 1.8%; Score 29; DB 7; Length 1267;

Best Local Similarity 100.0%; Pred. No. 0.089;

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1574 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 1602

Db 1222 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 1250

RESULT 10

ABL86191

ID ABL86191 standard; cDNA; 307 BP.

AC ABL86191;

XX 17-MAY-2002 (first entry)

XX DE Human ovarian cancer related cDNA clone SEQ ID NO:9169.

XX Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.

XX Homo sapiens.

XX WO200192581-A2.

XX 06-DEC-2001.

XX 29-MAY-2001; 2001WO-US017756.

XX 26-MAY-2000; 2000US-0207484P.

XX PA (CORI-) CORIXA CORP.  
 XX PI Algate PA, Harlocker SL, Jones R;  
 XX DR WPI; 2002-122075/16.  
 XX DR  
 XX PT Composition for therapy and diagnosis of ovarian cancer comprising  
 XX PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding  
 XX PT polypeptide, antibody specific to polypeptide or T cell expressing  
 XX PT polypeptide.  
 XX PS Claim 1; SEQ ID NO 9169; 489pp; English.  
 XX CC The present invention describes a composition (I) comprising: carriers  
 XX CC and immunostimulants; and a polypeptide (II) of a ovarian tumor  
 XX CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1)  
 XX CC from the 10912 nucleotide sequences as given in AB17023 to AB187934,  
 XX CC (III) encoding (II) having a sequence (S2), a T cell population of (II),  
 XX CC or antigen presenting cells that express (II). (I) has cytostatic  
 XX CC activity. An oligonucleotide (IV) that hybridises to (S1) can be used for  
 XX CC detecting ovarian cancer in a patient's biological sample preferably  
 XX CC serum or ovarian tissue. The method comprises contacting a biological  
 XX CC sample from a patient with (IV), detecting the amount of polynucleotide  
 XX CC hybridising to (IV) and comparing the amount to a predetermined cutoff  
 XX CC value and thereby detecting ovarian cancer in the patient, where the  
 XX CC amount of polynucleotide hybridising to (IV) is detected preferably by  
 XX CC polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is  
 XX CC useful for stimulating and/or expanding T cells specific for an ovarian  
 XX CC tumor protein comprising contacting T cells with (III) or (II). (III) is  
 XX CC useful in design and preparation of ribozyme molecules for inhibiting  
 XX CC expression of the tumor polypeptides and proteins in tumour cells; and  
 XX CC to isolate a full length gene from a suitable library e.g., a tumour cDNA  
 XX CC library using well known techniques  
 XX SQ Sequence 307 BP; 128 A; 50 C; 62 G; 67 T; 0 U; 0 Other;  
 Query Match 1.7%; Score 28; DB 6; Length 307;  
 Best Local Similarity 100.0%; Pred. No. 0.27; Mismatches 0; Indels 0; Gaps 0;  
 Matches 28; Conservative 0;  
 QY 1575 TTATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1602  
 DB 265 TTATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 292  
 RESULT 11  
 AAS60367/c  
 ID AAS60367 standard; cDNA; 339 BP.  
 AC AAS60367;  
 DT 29-JAN-2002 (first entry)  
 DE Human cancer agent-resistance marker #241.  
 KW Human; cancer cell marker; TAXOL; cytostatic; tumour; carcinoma;  
 KW squamous cell carcinoma; sarcoma; fibrosarcoma; leukaemia;  
 KW lymphocytic leukaemia; lymphoma; plasmocytoma; reticulum cell sarcoma;  
 KW Hodgkin's disease; glioma; sb.  
 OS Homo sapiens.  
 PN WO200179556-A2.  
 PD 25-OCT-2001.  
 PF 13-APR-2001; 2001WO-US012132.  
 XX 1A-APR-2000; 2000US-Q197538P.  
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 XX PT Evaluating the toxicity of an agent; useful in drug development or in  
 XX PT determining toxicological responses to a new drug by determining the  
 XX PT expression of rat toxicologically relevant genes in the test animal in

PI Lillie J, Brown JL, Bolt A, Van Huffel C;  
 XX WPI; 2001-602933/68.  
 XX Novel nucleic acid, used as a marker to determine the effectiveness of  
 XX PT using TAXOL to treat cancer cell growth in individuals.  
 XX PS Claim 1; Page 199; 527pp; English.  
 XX CC The invention relates to 1046 novel nucleic acids which are used as  
 XX CC markers for determining the sensitivity of a cancer cell to the  
 XX CC anticancer agent TAXOL. Cancer cells can be treated with TAXOL when they  
 XX CC are shown to express one of the 242 sensitivity markers or the cells are  
 XX CC shown not to express one of the 804 resistance markers. The methods can  
 XX CC be used to determine the effectiveness of TAXOL in the treatment of  
 XX CC cancer cell growth in an individual. The markers can be used as targets  
 XX CC in developing anti-cancer agents such as chemotherapeutic compounds. The  
 XX CC markers can also be used as targets in developing treatments for cancer,  
 XX CC particularly those cancers which display resistance to agents and exhibit  
 XX CC expression of the markers. The anticancer agents developed by the novel  
 XX CC method can be used to treat cancer. Probes based on the markers can be  
 XX CC used to detect transcripts or genomic sequences corresponding to the  
 XX CC markers, in the identification of cells or tissues which mis-express the  
 XX CC protein. Cancers which may be targeted include carcinoma (e.g. squamous  
 XX CC cell carcinoma), sarcoma (e.g. fibrosarcoma) leukaemia (e.g. lymphocytic  
 XX CC leukaemia), lymphoma, plasmocytoma, reticulum cell sarcoma, Hodgkin's  
 XX CC disease and tumours (e.g. glioma). The present sequence is one of the  
 XX CC 1046 novel cancer cell markers  
 XX SQ Sequence 339 BP; 110 A; 73 C; 64 G; 90 T; 0 U; 2 Other;  
 Query Match 1.7%; Score 28; DB 4; Length 339;  
 Best Local Similarity 100.0%; Pred. No. 0.27; Mismatches 0; Indels 0; Gaps 0;  
 Matches 28; Conservative 0;  
 QY 1573 ATTATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1600  
 DB 31 ATTATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4  
 RESULT 12  
 ABT09151/c  
 ID ABT09151 standard; DNA; 419 BP.  
 AC ABT09151;  
 DT 05-DEC-2002 (first entry)  
 DE Phase-1 Rat CT gene SEQ ID No 239.  
 DE Rat; toxicity study; rat toxic response gene; toxicological response;  
 KW drug development; phase-1 rat CT gene; db.  
 OS Rattus sp.  
 XX WO200266682-A2.  
 PD 29-AUG-2002.  
 PF 29-JAN-2002; 2002WO-US002935.  
 XX 29-JAN-2001; 2001US-0264933P.  
 PR 26-JUL-2001; 2001US-0308161P.  
 XX (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.  
 PI Farris G, Hicken SH, Farr SB;  
 DR WPI; 2002-674961/72.  
 XX Evaluating the toxicity of an agent; useful in drug development or in  
 XX PT determining toxicological responses to a new drug by determining the  
 XX PT expression of rat toxicologically relevant genes in the test animal in

PT response to the test agent.

PS Disclosure, Page 165; 388pp; English.

XX The invention relates to a method used for evaluating the toxicity of an agent comprising determining the expression of a rat toxic response gene(s) in the test animal in response to the agent. The method is useful in drug development, particularly for conducting toxicity studies and analysis before a new drug or compound is approved for human consumption or use. The method is also useful in determining toxicological responses to a new drug. This polynucleotide sequence represents a phase-1 rat CT gene of the invention

XX Sequence 419 BP; 153 A; 89 C; 85 G; 92 T; 0 U; 0 Other;

SEQ

Query Match 1.7%; Score 28; DB 6; Length 419;  
Best Local Similarity 100.0%; Pred. No. 0.26;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1574 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 1601  
|||||  
Db 28 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 1

RESULT 13  
ABA57560/c  
ID ABA57560 standard; DNA; 450 BP.  
XX  
AC ABA57560;  
XX  
XX 01-FEB-2002 (first entry)  
XX  
XX Human foetal liver single exon nucleic acid probe #5865.  
XX Human; foetal liver; Gene expression; single exon nucleic acid probe; ss.  
XX Homo sapiens.  
XX WO200157277-A2.  
XX  
XX 09-AUG-2001.  
XX  
XX 30-JAN-2001; 2001WO-US000669.  
XX  
XX 04-FEB-2000; 2000US-0180312P.  
XX 26-MAY-2000; 2000US-0207456P.  
XX 30-JUN-2000; 2000US-00608408.  
XX 03-AUG-2000; 2000US-00632366.  
XX 21-SEP-2000; 2000US-0234687P.  
XX 27-SEP-2000; 2000US-0236359P.  
XX 04-OCT-2000; 2000GB-00024263.  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-483447/52.  
XX Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human fetal liver.  
XX Claim 1; SEQ ID NO 5865; 639pp + Sequence Listing; English.  
XX The invention relates to a single exon nucleic acid probe for measuring human gene expression in a sample derived from human foetal liver. The single exon nucleic acid probes may be used for predicting, measuring and displaying gene expression in samples derived from human fetal liver. The present sequence is a single exon nucleic acid probe of the invention.  
XX Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/published\_sequences  
XX Sequence 450 BP; 123 A; 82 C; 95 G; 150 T; 0 U; 0 Other;

Query Match 1.7%; Score 28; DB 4; Length 450;  
Best Local Similarity 100.0%; Pred. No. 0.26;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1575 TTATGCAAAAAAAAAAAAAAAAAAAAAA 1602  
|||||  
Db 85 TTATGCAAAAAAAAAAAAAAAAAAAAAA 58

RESULT 14  
AAI37117/c  
ID AAI37117 standard; DNA; 450 BP.  
XX  
XX AAI37117;  
XX  
XX 17-OCT-2001 (first entry)  
XX  
XX Probe #5803 used to measure gene expression in human placenta sample.  
XX Probe; microarray; human; placenta; antenatal diagnosis;  
XX genetic disorder; ss.  
XX Homo sapiens.  
XX  
XX WO200157272-A2.  
XX  
XX 09-AUG-2001.  
XX  
XX 30-JAN-2001; 2001WO-US000663.  
XX  
XX 04-FEB-2000; 2000US-0180312P.  
XX 26-MAY-2000; 2000US-0207456P.  
XX 30-JUN-2000; 2000US-00608408.  
XX 03-AUG-2000; 2000US-00632366.  
XX 21-SEP-2000; 2000US-0234687P.  
XX 27-SEP-2000; 2000US-0236359P.  
XX 04-OCT-2000; 2000GB-00024263.  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-488897/53.  
XX Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human placenta.  
XX Claim 25; SEQ ID NO 5803; 654pp; English.  
XX The present invention relates to single exon nucleic acid probes (SENP). The present sequence is one such probe. The probes are useful for producing a microarray for predicting, measuring and displaying gene expression in samples derived from human placenta. The probes are useful for antenatal diagnosis of human genetic disorders  
XX Sequence 450 BP; 123 A; 82 C; 95 G; 150 T; 0 U; 0 Other;

Query Match 1.7%; Score 28; DB 4; Length 450;  
Best Local Similarity 100.0%; Pred. No. 0.26;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1575 TTATGCAAAAAAAAAAAAAAAAAAAAAA 1602  
|||||  
Db 85 TTATGCAAAAAAAAAAAAAAAAAAAAAA 58

RESULT 15  
ABA27022/c  
ID ABA27022 standard; DNA; 450 BP.  
XX  
XX ABA27022;  
XX

DT 23-JAN-2002 (first entry)  
XX Probe #5488 for gene expression analysis in human heart cell sample.  
DE Human; gene expression; heart; microarray; vascular system; probe;  
XX cardiovascular disease; hypertension; cardiac arrhythmia;  
KW congenital heart disease; ss.  
XX  
OS Homo sapiens.  
XX WO200157274-A2.  
XX  
XX 09-AUG-2001.  
XX  
XX 30-JAN-2001; 2001WO-US000666.  
XX  
XX 04-FEB-2000; 2000US-0180312P.  
XX 26-MAY-2000; 2000US-0207456P.  
XX 30-JUN-2000; 2000US-00608408.  
XX 03-AUG-2000; 2000US-00632366.  
XX 21-SEP-2000; 2000US-0234687P.  
XX 27-SEP-2000; 2000US-0236359P.  
XX 04-OCT-2000; 2000GB-00024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-488899/53.  
XX  
XX Single exon nucleic acid probes for analyzing gene expression in human  
PT hearts.  
XX  
XX Claim 1; SEQ ID NO 5488; 530pp; English.  
XX  
XX The present invention relates to single exon nucleic acid probes for  
CC measuring human gene expression in a sample derived from human heart. The  
CC present sequence is one such probe. The probes may be used for  
CC predicting, measuring and displaying gene expression in samples derived  
CC from the human heart via microarrays. By measuring gene expression, the  
CC probes are useful for predicting, diagnosing, grading, staging,  
CC monitoring and prognosing diseases of the human heart and vascular system  
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and  
CC congenital heart disease. Note: The sequence data for this patent did not  
CC form part of the printed specification, but was obtained in electronic  
CC format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 450 BP; 123 A; 82 C; 95 G; 150 T; 0 U; 0 Other;  
XX  
Query Match 1.7%; Score 28; DB 4; Length 450;  
Best Local Similarity 100.0%; Pred. No. 0.26;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1575 TTATCGAAAAAAAAAAAAAAAAAAAAA 1602  
DB 85 TTATCGAAAAAAAAAAAAAAAAAAAAA 58

Search completed: August 2, 2004, 11:24:58  
Job time : 672 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2004, 08:46:58 ; Search time 176 Seconds

(without alignments)

5051.314 Million cell updates/sec

**Title:** US-10-031-331B-39

**Perfect score:**

Sequence: 1 cacaccgttgatttaaccat.....aaaaaaaaaaaaaaaaaaaaa 1602

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA: \*

```

1: /cgn2_6/prodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/prodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/prodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/prodata/2/ina/PCTR_COMB.seq.*
6: /cgn2_6/prodata/2/ina/backfiles1.seq.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query		DB	ID	Description	
		Match	Length				
1	27	1.7	233	4	US-09-621-976-10675	Sequence 10675, A	
2	27	1.7	1319	4	US-09-643-067-17	Sequence 17, Appl	
3	26	1.6	32	4	US-09-619-103-14	Sequence 2, Appl	
4	26	1.6	38	4	US-09-619-103-2	Sequence 2, Appl	
5	26	1.6	249	4	US-09-621-976-19144	Sequence 19144, A	
6	26	1.6	787	4	US-09-621-976-1878	Sequence 1878, Ap	
7	26	1.6	971	3	US-09-248-335-65	Sequence 65, Appl	
8	26	1.6	1181	4	US-09-149-476-310	Sequence 310, App	
9	26	1.6	1195	1	US-08-373-858-1	Sequence 1, Appl	
10	26	1.6	1195	1	US-08-342-7868-1	Sequence 1, Appl	
11	26	1.6	1212	4	US-09-149-476-186	Sequence 186, App	
12	26	1.6	1359	3	US-09-387-574-11	Sequence 11, Appl	
13	26	1.6	1359	4	US-08-668-096-11	Sequence 11, Appl	
14	26	1.6	1334	1	US-08-300-303A-6	Sequence 6, Appl	
15	26	1.6	1534	4	US-08-988-197-6	Sequence 6, Appl	
16	26	1.6	1946	4	US-09-882-835-1	Sequence 1, Appl	
17	26	1.6	2971	4	US-09-482-273-37	Sequence 37, Appl	
18	26	1.6	3275	4	US-09-370-838-151	Sequence 151, App	
c	20	26	1.6	44848	US-09-435-739-42	Sequence 42, Appl	
	19	26	1.6	90050	US-09-245-041-5	Sequence 5, Appl	
21	25	1.6	57	1	US-08-120-827-93	Sequence 93, Appl	
22	25	1.6	57	1	US-08-478-675-93	Sequence 93, Appl	
23	25	1.6	57	3	US-08-199-737-53	Sequence 59, Appl	
c	24	25	1.6	57	4	US-09-056-933A-59	Sequence 59, Appl
	25	25	1.6	104	4	US-09-621-976-11198	Sequence 11198, A
26	25	1.6	114	4	US-09-621-976-9602	Sequence 9602, Ap	
27	25	1.6	173	4	US-09-621-976-18333	Sequence 18333, A	

## ALIGNMENTS

## RESULT 1

```

US-09-621-976-10675
; Sequence 10675, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Ests and Encoded H
; FILE REFERENCE: GNSST 054PR2
; CURRENT APPLICATION NUMBER: US/09/621,9
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 10675
; LENGTH: 233
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-10675

```

Query Match 1.7%; Score 27; DB 4; Length 233;  
Best Local Similarity 100.0%; Pred. No. 0.008;  
Matches 27; Conservative 0; Mismatches 0; Indels

Qy	1576	TATGC	AAAAAAAAAAAAAAAAAAAA	1602
Db	178	TATGC	AAAAAAAAAAAAAAAAAAAA	204

## RESULT 2

```

US-09-443-067-17
: Sequence 17, Application US/09443067
: Patent No. 6627794
: GENERAL INFORMATION:
: APPLICANT: COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH
: APPLICANT: ORGANISATION
: TITLE OF INVENTION: Polyphenol oxidase genes from banana, 1
: TITLE OF INVENTION: pineapple
: FILE REFERENCE:
: CURRENT APPLICATION NUMBER: US/09/443,067
: CURRENT FILING DATE: 1999-11-18
: EARLIER APPLICATION NUMBER: US 08/976, 222
: EARLIER FILING DATE: 1997-11-21
: EARLIER APPLICATION NUMBER: PCT/AU98/00362
: EARLIER FILING DATE: 1998-05-19
: EARLIER APPLICATION NUMBER: AU P8398
: EARLIER FILING DATE: 1995-05-23
: EARLIER APPLICATION NUMBER: AU P96849
: EARLIER FILING DATE: 1997-05-19
: EARLIER APPLICATION NUMBER: AU P95600

```

```
; EARLIER FILING DATE: 1995-09-26
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 17
; LENGTH: 1319
; TYPE: DNA
; ORGANISM: pineapple
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1053)
US-09-443-067-17

Query Match      1.7%; Score 27; DB 4; Length 1319;
Best Local Similarity 100.0%; Pred. No. 0.0076;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1576 TATCCAAAAA..... 1602
DB 1288 TATCCAAAAA..... 1314

RESULT 3
US-09-619-103-14
; Sequence 14, Application US/09619103
; Patent No. 6429300
; GENERAL INFORMATION:
; APPLICANT: Kurz, Markus
; APPLICANT: Lohse, Peter
; APPLICANT: Wagner, Richard
; TITLE OF INVENTION: Peptide Acceptor Ligation Methods
; FILE REFERENCE: 50036/031002
; CURRENT APPLICATION NUMBER: US/09/619,103
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 60/145,834
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 32
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: designed sequence to act as a linker
US-09-619-103-14

Query Match      1.6%; Score 26; DB 4; Length 32;
Best Local Similarity 96.2%; Pred. No. 0.024;
Matches 25; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1577 ATGCA..... 1602
DB 7 AUGCA..... 32

RESULT 4
US-09-619-103-2
; Sequence 2, Application US/09619103
; Patent No. 6429300
; GENERAL INFORMATION:
; APPLICANT: Kurz, Markus
; APPLICANT: Lohse, Peter
; APPLICANT: Wagner, Richard
; TITLE OF INVENTION: Peptide Acceptor Ligation Methods
; FILE REFERENCE: 50036/031002
; CURRENT APPLICATION NUMBER: US/09/619,103
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 60/145,834
; PRIOR FILING DATE: 1999-07-27
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 38
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: designed sequence to act as a linker
US-09-619-103-2

Query Match      1.6%; Score 26; DB 4; Length 38;
Best Local Similarity 100.0%; Pred. No. 0.024;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1577 ATGCA..... 1602
DB 6 ATGCA..... 31

RESULT 5
US-09-621-976-19144
; Sequence 19144, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 19144
; LENGTH: 249
; TYPE: DNA
; ORGANISM: Homo sapiens
; OTHER INFORMATION:
US-09-621-976-19144

Query Match      1.6%; Score 26; DB 4; Length 249;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1577 ATGCA..... 1602
DB 188 ATGCA..... 213

RESULT 6
US-09-621-976-1878
; Sequence 1878, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 1878
; LENGTH: 787
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 152..673
; NAME/KEY: sig.peptide
; LOCATION: 152..208
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 4.09999990463257
; OTHER INFORMATION: seq LLIGQRCSLKVG/QE
US-09-621-976-1878

Query Match      1.6%; Score 26; DB 4; Length 787;
Best Local Similarity 100.0%; Pred. No. 0.022;
```

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1577 ATGCAAAAAAAAAAAAAAAAAAAAAA 1602  
DB 699 ATGCAAAAAAAAAAAAAAAAAAAAAA 724

## RESULT 7

US-09-248-335-65  
; Sequence 65, Application US/09248335  
; Patent No. 6096504  
; GENERAL INFORMATION:  
; APPLICANT: MCGONIGLE, BRIAN  
; TITLE OF INVENTION: PLANT GLUTATHIONE-S-TRANSFERASE ENZYMES  
; FILE REFERENCE: CL-1128-A  
; CURRENT APPLICATION NUMBER: US/09/248,335  
; CURRENT FILING DATE: 1999-02-10  
; EARLIER APPLICATION NUMBER: 08/924,759  
; EARLIER FILING DATE: 1997-September-05  
; NUMBER OF SEQ ID NOS: 74  
; SOFTWARE: Microsoft Word Version 7.0A  
; SEQ ID NO 65  
; LENGTH: 971  
; TYPE: DNA  
; ORGANISM: maize  
US-09-248-335-65

Query Match 1.6%; Score 26; DB 3; Length 971;  
Best Local Similarity 100.0%; Pred. No. 0.022;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1577 ATGCAAAAAAAAAAAAAAAAAAAAAA 1602  
DB 940 ATGCAAAAAAAAAAAAAAAAAAAAAA 965

## RESULT 8

US-09-149-476-310  
; Sequence 310, Application US/09149476  
; Patent No. 6420526  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 186 Human Secreted proteins  
; FILE REFERENCE: P2002P1  
; CURRENT APPLICATION NUMBER: US/09/149,476  
; CURRENT FILING DATE: 1998-09-08  
; EARLIER APPLICATION NUMBER: PCT/US98/04493  
; EARLIER FILING DATE: 1998-03-06  
; EARLIER APPLICATION NUMBER: 60/040,162  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,333  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/038,621  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,626  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,334  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,336  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,163  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/047,600  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,615  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,597  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,502  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,633  
; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,583  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,617  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,618  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,503  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,592  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,581  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,584  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,500  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,587  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,492  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,598  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,613  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,582  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,596  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,612  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,632  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,601  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/043,580  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,568  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,314  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,569  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,311  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,671  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,674  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,669  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,312  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,313  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,672  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,315  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/048,974  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/056,886  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,877  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,889  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,893  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,630  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,878  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,662



EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,872  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,882  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,887  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,908  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/048,964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/057,650  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/056,884  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/057,669  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/049,610  
EARLIER FILING DATE: 1997-06-13  
EARLIER APPLICATION NUMBER: 60/061,060  
EARLIER FILING DATE: 1997-10-02

EARLIER APPLICATION NUMBER: 60/056,875  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,862  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,887  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,908  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/048,964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/057,650  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/056,884  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/057,669  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/049,610  
EARLIER FILING DATE: 1997-06-13  
EARLIER APPLICATION NUMBER: 60/061,060  
EARLIER FILING DATE: 1997-10-02

Query Match 1.6%; Score 26; DB 4; Length 1181;  
Best Local Similarity 100.0%; Pred. No. 0.022;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1577 ATGCACAAAAA 1602  
DB 1123 ATGCACAAAAA 1148

RESULT 9  
US-08-373-858-1  
Sequence 1, Application US/08373858  
Patent No. 5633155  
GENERAL INFORMATION:  
APPLICANT: Kim, Man-Keun  
APPLICANT: Lee, Kwan-Ho  
APPLICANT: Na, Byeong-Kook  
APPLICANT: Jeong, Han-Seung  
APPLICANT: Choo, Kyu-Whan  
APPLICANT: Moon, Young-Ho  
APPLICANT: Jeon, Hong-Seob  
TITLE OF INVENTION: Expression Vector for Phytoacta  
TITLE OF INVENTION: Antiviral Protein and Process for Preparing Transgenic  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESS: Darby & Darby  
STREET: 805 Third Avenue  
CITY: New York  
STATE: New York  
COUNTRY: US  
ZIP: 10022-7513  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/373,858  
FILING DATE: 18-JAN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Ludwig, Peter  
REGISTRATION NUMBER: 25,351  
REFERENCE/DOCKET NUMBER: 0136/17986-US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-527-7770  
TELEFAX: 212-753-6237  
TELEX: 236687  
INFORMATION FOR SEQ ID NO. 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1195 base pairs

TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Phytolacca americana  
TISSUE TYPE: Leaf  
US-08-373-858-1

Query Match 1.6%; Score 26; DB 1; Length 1195;  
Best Local Similarity 100.0%; Pred. No. 0.022;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1577 ATGCAAAAAAAAAAAAAAAAAAAAAA 1602  
DB 1169 ATGCAAAAAAAAAAAAAAAAAAAAAA 1194

## RESULT 10

US-08-342-786B-1  
Sequence 1, Application US/08342786B  
Patent No. 5648234

GENERAL INFORMATION:  
APPLICANT: Moon, Young-Ho

APPLICANT: Jeon, Hong-Seob

APPLICANT: Choi, Kyu-Whan

APPLICANT: Lee, Kwan-Ho

APPLICANT: Kim, Man-Keun

TITLE OF INVENTION: A No. 5648234el Expression Vector for Phytolacca

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Darby & Darby

STREET: 805 Third Ave.

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/342,786B

FILING DATE: 21-NOV-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/147,024

FILING DATE: 01-NOV-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Ludwig, S. Peter

REGISTRATION NUMBER: 25,351

REFERENCE/DOCKET NUMBER: 0136/08862-US1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212)527-7700

TELEFAX: (212)753-6237

TELEX: 236687

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1195 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Phytolacca americana L.

IMMEDIATE SOURCE:

CLONE: PAP

## US-08-342-786B-1

Query Match 1.6%; Score 26; DB 1; Length 1195;  
Best Local Similarity 100.0%; Pred. No. 0.022;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1577 ATGCAAAAAAAAAAAAAAAAAAAAAA 1602  
DB 1169 ATGCAAAAAAAAAAAAAAAAAAAAAA 1194

## RESULT 11

US-08-149-476-186

Sequence 186, Application US/09149476

Patent No. 8420526

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: 186 Human Secreted proteins

FILE REFERENCE: P2002P1

CURRENT APPLICATION NUMBER: US/09/149,476

EARLIER APPLICATION NUMBER: PCT/US98/04493

EARLIER FILING DATE: 1998-09-08

EARLIER FILING DATE: 1998-03-06

EARLIER APPLICATION NUMBER: 60/040,162

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,333

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/038,621

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,626

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,334

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,336

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,163

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/047,600

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,615

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,597

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,502

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,633

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,583

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,617

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,618

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,503

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,592

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,581

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,584

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,500

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,587

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,492

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,598

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,613

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,582

EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,596  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,612  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,632  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,601  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/043,580  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,568  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,314  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,569  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,311  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,671  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,674  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,669  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,312  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,313  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,672  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,315  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/048,974  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/056,886  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,877  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,889  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,893  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,630  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,878  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,662  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,872  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,882  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,637  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,903  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,888  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,879  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,880  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,894  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,911  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,636  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,874  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,910  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,864

; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,631  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,845  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,892  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/057,761  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/047,595  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,599  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,588  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,585  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,586  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,590  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,594  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,589  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,593  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,614  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/043,578  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,576  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/047,501  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/043,670  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/056,632  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,664  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,876  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,881  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,909  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,875  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,862  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,887  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,908  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/048,964  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/057,650  
; EARLIER FILING DATE: 1997-09-05  
; EARLIER APPLICATION NUMBER: 60/056,884  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/057,669  
; EARLIER FILING DATE: 1997-09-05  
; EARLIER APPLICATION NUMBER: 60/049,610  
; EARLIER FILING DATE: 1997-06-13  
; EARLIER APPLICATION NUMBER: 60/061,060  
; EARLIER FILING DATE: 1997-10-02

Query Match 1.68; Score 26; DB 4; Length 1212;  
Best Local Similarity 100.0%; Pred. No. 0.022;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1577 ATGCAAAAAAAAAAAAAAAAAAAAAA 1602

```
Db 1154 ATGCACAAAAAAAAAAAAAAAAAAAA 1179
|||||
RESULT 12
US-09-387-574-11
; Sequence 11, Application US/09387574
; Patent No. 6168951
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Kinney, Tony
; APPLICANT: Rafalski, Antoni
; TITLE OF INVENTION: Plant Geranylgeranyl Transferases
; FILE REFERENCE: BB-1239
; CURRENT APPLICATION NUMBER: US/09/387,574
; EARLIER FILING DATE: 1999-08-31
; EARLIER APPLICATION NUMBER: 60/098,743
; EARLIER FILING DATE: September 1, 1998
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 11
; LENGTH: 1359
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-387-574-11
Query Match 1.6%; Score 26; DB 3; Length 1359;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1577 ATGCACAAAAAAAAAAAAAAAAAAAA 1602
DB 1289 ATGCACAAAAAAAAAAAAAAAAAAAA 1314
|||||
RESULT 13
US-09-668-096-11
; Sequence 11, Application US/09668096
; Patent No. 6312954
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Kinney, Tony
; APPLICANT: Rafalski, Antoni
; TITLE OF INVENTION: Plant Geranylgeranyl Transferases
; FILE REFERENCE: BB1239 US NA DIV
; CURRENT APPLICATION NUMBER: US/09/668,096
; EARLIER FILING DATE: 2000-09-22
; EARLIER APPLICATION NUMBER: 60/098,743
; EARLIER FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 09/387,534
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 11
; LENGTH: 1359
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-668-096-11
Query Match 1.6%; Score 26; DB 4; Length 1359;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1577 ATGCACAAAAAAAAAAAAAAAAAAAA 1602
DB 1289 ATGCACAAAAAAAAAAAAAAAAAAAA 1314
|||||
RESULT 14
US-08-300-903A-6
; Sequence 6, Application US/08300903A
; Patent No. 5591630
; GENERAL INFORMATION:
; APPLICANT: Anderson, Dirk M
; APPLICANT: Giri, Judith G
; TITLE OF INVENTION: Interleukin-15 Receptors
; CORRESPONDENCE ADDRESS:
; ADDRESS: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: Apple Macintosh
; SOFTWARE: Microsoft Word for Apple, Version 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/300,903A
; FILING DATE: 06-SEPTEMBER-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/236,919
; FILING DATE: 06-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,695
; REFERENCE/DOCKET NUMBER: 2822-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0430
; TELEFAX: 206-233-0644
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1534 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..753
US-08-300-903A-6
Query Match 1.6%; Score 26; DB 1; Length 1534;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1577 ATGCACAAAAAAAAAAAAAAAAAAAA 1602
DB 1476 ATGCACAAAAAAAAAAAAAAAAAAAA 1501
|||||
RESULT 15
US-08-988-197-6
; Sequence 6, Application US/08988197
; Patent No. 6548065
; GENERAL INFORMATION:
; APPLICANT: Anderson, Dirk M
; APPLICANT: Giri, Judith G
; TITLE OF INVENTION: Interleukin-15 Receptors
; CORRESPONDENCE ADDRESS:
; ADDRESS: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
```

OPERATING SYSTEM: Apple Operating System 7.1  
SOFTWARE: Microsoft Word for Apple, Version 5.1a  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/988,197  
FILING DATE:  
CLASSIFICATION: 121097  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/300,903  
FILING DATE: 06-SEPTEMBER-1994  
APPLICATION NUMBER: USSN 08/236,919  
FILING DATE: 06-MAY-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,695  
REFERENCE/DOCKET NUMBER: 2822-A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206-587-0430  
TELEFAX: 206-233-0644  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1534 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..753  
US-08-988-197-6

Query Match 1.6%; Score 26; DB 4; Length 1534;  
Best Local Similarity 100.0%; Pred. No. 0.022;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1577 ATGCAAAAAAAAAAAAAAAAAAAAAA 1602  
Db 1476 ATGCAAAAAAAAAAAAAAAAAAAAAA 1501

Search completed: August 2, 2004, 14:51:07  
Job time : 177 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2004, 13:13:14 ; Search time 767 Seconds  
(without alignments)  
10240.978 Million cell updates/sec

Title: US-10-031-331b-39  
Perfect score: 1602  
Sequence: 1 cacacgttgattaccat.....aaaaaaaaaaaaaaaa 1602

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Capext 60.0

Searched: 3222919 seqs, 2451570024 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6445838

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Published Applications NA.\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	30	1.9	504	13	US-10-424-599-121946
C 2	30	1.9	523	17	US-10-021-323-11875
C 3	29	1.8	416	13	US-10-424-599-71220
C 4	29	1.8	483	10	US-09-770-961-873
C 5	29	1.8	496	17	US-10-021-323-4109
C 6	29	1.8	1267	9	US-09-764-864-317
C 7	29	1.8	43853	17	US-10-322-281-788
C 8	28	1.7	307	9	US-09-867-701-9169
C 9	28	1.7	339	9	US-09-834-975-368
C 10	28	1.7	450	9	US-09-864-761-5488
C 11	28	1.7	559	13	US-10-342-887-1047
C 12	28	1.7	559	13	US-10-172-118-1077
C 13	28	1.7	559	13	US-10-170-385-468
C 14	28	1.7	809	17	US-10-437-963-87473

C 15	28	1.7	1021	17	US-10-437-963-47572
C 16	28	1.7	1051	13	US-10-424-599-140317
C 17	28	1.7	1069	17	US-10-437-963-30744
C 18	28	1.7	2000	9	US-09-938-842A-4789
C 19	28	1.7	2000	11	US-09-938-842A-4789
C 20	28	1.7	2006	16	US-10-442-017-8
C 21	28	1.7	3183	15	US-10-000-897-29
C 22	28	1.7	45121	12	US-09-997-722-10
C 23	27	1.7	43	13	US-10-175-539A-45
C 24	27	1.7	159	13	US-10-085-783A-14143
C 25	27	1.7	159	16	US-10-242-535A-14143
C 26	27	1.7	424	15	US-10-198-846-13129
C 27	27	1.7	444	9	US-09-962-436-88
C 28	27	1.7	444	9	US-09-880-107-2880
C 29	27	1.7	444	9	US-09-967-768A-167
C 30	27	1.7	474	13	US-10-027-632-49689
C 31	27	1.7	474	16	US-10-027-632-49689
C 32	27	1.7	491	12	US-09-969-034-984
C 33	27	1.7	492	17	US-10-437-963-3082
C 34	27	1.7	505	13	US-10-424-599-45864
C 35	27	1.7	513	10	US-09-918-995-21477
C 36	27	1.7	542	13	US-10-027-632-5404
C 37	27	1.7	542	16	US-10-027-632-5404
C 38	27	1.7	681	15	US-10-086-543-976
C 39	27	1.7	854	15	US-10-198-846-6097
C 40	27	1.7	979	16	US-10-159-563-298
C 41	27	1.7	1164	17	US-10-437-963-28218
C 42	27	1.7	1170	16	US-10-264-049-927
C 43	27	1.7	1201	13	US-10-342-887-955
C 44	27	1.7	1201	13	US-10-172-118-955
C 45	27	1.7	1335	9	US-09-925-300-534

#### ALIGNMENTS

RESULT 1  
US-10-424-599-121946/c  
; Sequence 121946, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 121946  
; LENGTH: 504  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_81123C.1  
US-10-424-599-121946

Query Match 1.9%; Score 30; DB 13; Length 504;  
Best Local Similarity 100.0%; Pred. No. 0.00015;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1573 ATTTATGCAAAAAAAAAAAAAAAAAAAAAA 1602  
Db 44 ATTTATGCAAAAAAAAAAAAAAAAAAAAAA 15

RESULT 2  
US-10-021-323-11875  
; Sequence 11875, Application US/10021323  
; Publication No. US20040123340A1  
; GENERAL INFORMATION:

APPLICANT: Deikman, Jill  
APPLICANT: Feng, Paul C.C.  
APPLICANT: Fincher, Karen L.  
APPLICANT: Ziegler, Todd B.  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(52274)B  
CURRENT APPLICATION NUMBER: US/10/021.323  
CURRENT FILING DATE: 2001-12-12  
PRIOR APPLICATION NUMBER: US 60/255, 619  
PRIOR FILING DATE: 2000-12-14  
NUMBER OF SEQ ID NOS: 17880  
SEQ ID NO 11875  
LENGTH: 523  
TYPE: DNA  
ORGANISM: Gossypium hirsutum  
FEATURE:  
OTHER INFORMATION: Clone ID: LIB3829-014-Q6-K6-A4  
US-10-021-323-11875

Query Match 1.9%; Score 30; DB 17; Length 523;  
Best Local Similarity 100.0%; Pred. No. 0.00015;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1573 ATTATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1602  
Db 307 ATTATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 336

RESULT 3  
US-10-424-599-71220  
Sequence 71220, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 71220  
LENGTH: 416  
TYPE: DNA  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT MRT3847\_35323C.1  
US-10-424-599-71220

Query Match 1.8%; Score 29; DB 13; Length 416;  
Best Local Similarity 100.0%; Pred. No. 0.00049;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1574 TTATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1602  
Db 144 TTATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 172

RESULT 4  
US-09-770-961-873/c  
Sequence 873, Application US/09770961  
Publication No. US20030115639A1  
GENERAL INFORMATION:  
APPLICANT: Gorlach, Jorn  
APPLICANT: An, Yong-Giang  
APPLICANT: Hamilton, Carol M.  
APPLICANT: Price, Jennifer L.  
APPLICANT: Raines, Tracy M.  
APPLICANT: Yu, Yang  
APPLICANT: Rameaka, Joshua G.

APPLICANT: Page, Amy  
APPLICANT: Matthew, Abraham V.  
APPLICANT: Ledford, Brooke L.  
APPLICANT: Woessner, Jeffrey P.  
APPLICANT: Haas, William David  
APPLICANT: Garcia, Carlos A.  
APPLICANT: Krickler, Naja  
APPLICANT: Slader, Ted  
APPLICANT: Davis, Keith R.  
APPLICANT: Allen, Keith  
APPLICANT: Hoffman, Neil  
APPLICANT: Hurlban, Patrick  
TITLE OF INVENTION: Expressed Sequences of Arabidopsis  
FILE REFERENCE: 2026 (PARA-015PRV)  
CURRENT APPLICATION NUMBER: US/09/770,961  
CURRENT FILING DATE: 2001-01-26  
PRIOR APPLICATION NUMBER: 60/178,466  
PRIOR FILING DATE: 2000-01-27  
NUMBER OF SEQ ID NOS: 999  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 873  
LENGTH: 483  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
US-09-770-961-873

Query Match 1.8%; Score 29; DB 10; Length 483;  
Best Local Similarity 100.0%; Pred. No. 0.0005;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1574 TTATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1602  
Db 48 TTATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 20

RESULT 5  
US-10-021-323-4109  
Sequence 4109, Application US/10021323  
Publication No. US20040123340A1  
GENERAL INFORMATION:  
APPLICANT: Deikman, Jill  
APPLICANT: Feng, Paul C.C.  
APPLICANT: Fincher, Karen L.  
APPLICANT: Ziegler, Todd B.  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(52274)B  
CURRENT APPLICATION NUMBER: US/10/021,323  
CURRENT FILING DATE: 2001-12-12  
PRIOR APPLICATION NUMBER: US 60/255, 619  
PRIOR FILING DATE: 2000-12-14  
NUMBER OF SEQ ID NOS: 17880  
SEQ ID NO 4109  
LENGTH: 496  
TYPE: DNA  
ORGANISM: Gossypium hirsutum  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)-(496)  
OTHER INFORMATION: unsure at all n locations  
OTHER INFORMATION: Clone ID: LIB3825-031-Q6-K6-B6  
US-10-021-323-4109

Query Match 1.8%; Score 29; DB 17; Length 496;  
Best Local Similarity 100.0%; Pred. No. 0.0005;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1574 TTATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1602  
Db 194 TTATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 222

APPLICANT: Deikman, Jill  
APPLICANT: Feng, Paul C.C.  
APPLICANT: Fincher, Karen L.  
APPLICANT: Ziegler, Todd B.  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(52274)B  
CURRENT APPLICATION NUMBER: US/10/021.323  
CURRENT FILING DATE: 2001-12-12  
PRIOR APPLICATION NUMBER: US 60/255, 619  
PRIOR FILING DATE: 2000-12-14  
NUMBER OF SEQ ID NOS: 17880  
SEQ ID NO 11875  
LENGTH: 523  
TYPE: DNA  
ORGANISM: Gossypium hirsutum  
FEATURE:  
OTHER INFORMATION: Clone ID: LIB3829-014-Q6-K6-A4  
US-10-021-323-11875

Query Match 1.9%; Score 30; DB 17; Length 523;  
Best Local Similarity 100.0%; Pred. No. 0.00015;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1573 ATTATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1602  
Db 307 ATTATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 336

RESULT 3  
US-10-424-599-71220  
Sequence 71220, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 71220  
LENGTH: 416  
TYPE: DNA  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT MRT3847\_35323C.1  
US-10-424-599-71220

Query Match 1.8%; Score 29; DB 13; Length 416;  
Best Local Similarity 100.0%; Pred. No. 0.00049;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1574 TTATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1602  
Db 144 TTATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 172

RESULT 4  
US-09-770-961-873/c  
Sequence 873, Application US/09770961  
Publication No. US20030115639A1  
GENERAL INFORMATION:  
APPLICANT: Gorlach, Jorn  
APPLICANT: An, Yong-Giang  
APPLICANT: Hamilton, Carol M.  
APPLICANT: Price, Jennifer L.  
APPLICANT: Raines, Tracy M.  
APPLICANT: Yu, Yang  
APPLICANT: Rameaka, Joshua G.



## RESULT 6

US-09-764-864-317  
; Sequence 317, Application US/09764864  
; Patent No. US20020132753A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: P1223  
; CURRENT APPLICATION NUMBER: US/09/764,864  
; PRIOR FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1792  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 317  
; LENGTH: 1267  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-864-317

Query Match 1.8%; Score 29; DB 9; Length 1267;  
Best Local Similarity 100.0%; Pred. No. 0.00052;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1574 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 1602  
Db 1222 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 1250

## RESULT 7

US-10-322-281-788/c  
; Sequence 788, Application US/10322281  
; Publication No. US20040126762A1  
; GENERAL INFORMATION:  
; APPLICANT: David W. Morris  
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer  
; FILE REFERENCE: 529452001000  
; CURRENT APPLICATION NUMBER: US/10/322,281  
; CURRENT FILING DATE: 2002-12-17  
; NUMBER OF SEQ ID NOS: 866  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 788  
; LENGTH: 43853  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-322-281-788

Query Match 1.8%; Score 29; DB 17; Length 43853;  
Best Local Similarity 100.0%; Pred. No. 0.00061;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1574 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 1602  
Db 40399 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 40371

## RESULT 8

US-09-867-701-9169  
; Sequence 9169, Application US/09867701  
; Patent No. US20020132237A1  
; GENERAL INFORMATION:  
; APPLICANT: Aglate, Paul A.  
; APPLICANT: Jones, Robert  
; APPLICANT: Harlocker, Susan L.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.497  
; CURRENT APPLICATION NUMBER: US/09/867,701  
; CURRENT FILING DATE: 2001-05-29  
; NUMBER OF SEQ ID NOS: 10912  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9169  
; LENGTH: 307

; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-867-701-9169

Query Match 1.7%; Score 28; DB 9; Length 307;  
Best Local Similarity 100.0%; Pred. No. 0.0016;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1575 TTATGCAAAAAAAAAAAAAAAAAAAAAA 1602  
Db 265 TTATGCAAAAAAAAAAAAAAAAAAAAAA 292

## RESULT 9

US-09-834-975-368/c  
; Sequence 368, Application US/09834975  
; Patent No. US20020110815A1  
; GENERAL INFORMATION:  
; APPLICANT: Lillie, James  
; APPLICANT: Brown, Jeffrey  
; APPLICANT: Bolt, Andrew  
; APPLICANT: Van Hufel, Christophe  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS  
; TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY  
; FILE REFERENCE: MRI-016B  
; CURRENT APPLICATION NUMBER: US/09/834,975  
; CURRENT FILING DATE: 2001-04-13  
; PRIOR APPLICATION NUMBER: 60/197,538  
; PRIOR FILING DATE: 2000-04-14  
; NUMBER OF SEQ ID NOS: 1046  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 368  
; LENGTH: 339  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (1)...(339)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-834-975-368

Query Match 1.7%; Score 28; DB 9; Length 339;  
Best Local Similarity 100.0%; Pred. No. 0.0016;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1573 ATTTATGCAAAAAAAAAAAAAAAAAAAAAA 1600  
Db 31 ATTTATGCAAAAAAAAAAAAAAAAAAAAAA 4

## RESULT 10

US-09-864-761-5488/c  
; Sequence 5488, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aemica-x-1  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US/09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SEQ ID NO 5488  
LENGTH: 450  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC005479.2  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.5  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4  
US-09-864-761-5488

Query Match 1.7%; Score 28; DB 9; Length 450;  
Best Local Similarity 100.0%; Pred. No. 0.0016;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1575 TTATGCAAAAAAAAAAAAAAAAAAAAAA 1602  
DB 85 TTATGCAAAAAAAAAAAAAAAAAAAAAA 58

## RESULT 11

US-10-342-887-1077  
Sequence 1077, Application US/10342887.  
Publication No. US20040058340A1  
GENERAL INFORMATION:

APPLICANT: Dai, Hongyue  
APPLICANT: He, Yudong  
APPLICANT: Linsley, Peter S.  
APPLICANT: Mao, Mao  
APPLICANT: Roberts, Christopher J.  
APPLICANT: Van 't Veer, Laura Johanna  
APPLICANT: Van de Vijver, Marc J.  
APPLICANT: Bernards, Rene

TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients

FILE REFERENCE: 9301-188-999  
CURRENT APPLICATION NUMBER: US/10/342,897

PRIOR FILING DATE: 2003-01-15

PRIOR APPLICATION NUMBER: 60/298,918

PRIOR FILING DATE: 2001-06-18

PRIOR APPLICATION NUMBER: 60/380,710  
PRIOR FILING DATE: 2002-05-14  
PRIOR APPLICATION NUMBER: 10/172,118  
PRIOR FILING DATE: 2002-06-14  
NUMBER OF SEQ ID NOS: 2699  
SEQ ID NO 1077  
LENGTH: 559  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-342-887-1077

Query Match 1.7%; Score 28; DB 13; Length 559;  
Best Local Similarity 100.0%; Pred. No. 0.0016;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1575 TTATGCAAAAAAAAAAAAAAAAAAAAAA 1602  
DB 513 TTATGCAAAAAAAAAAAAAAAAAAAAAA 540

## RESULT 12

US-10-172-118-1077  
Sequence 1077, Application US/10172118  
Publication No. US20030224374A1  
GENERAL INFORMATION:

APPLICANT: Dai, Hongyue  
APPLICANT: He, Yudong  
APPLICANT: Linsley, Peter  
APPLICANT: Mao, Mao  
APPLICANT: Roberts, Chris  
APPLICANT: Van 't Veer, Laura  
APPLICANT: Van de Vijver, Marc  
APPLICANT: Bernards, Rene

TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients

FILE REFERENCE: 9301-175-999

CURRENT APPLICATION NUMBER: US/10/172,118

CURRENT FILING DATE: 2002-06-14

PRIOR APPLICATION NUMBER: 60/380,770

PRIOR FILING DATE: 2002-05-14

NUMBER OF SEQ ID NOS: 2699

SEQ ID NO 1077

LENGTH: 559

TYPE: DNA

ORGANISM: Homo sapiens

PUBLICATION INFORMATION:

DATABASE ACCESSION NUMBER: NM\_004708

DATABASE ENTRY DATE: 2001-06-18

US-10-172-118-1077

Query Match 1.7%; Score 28; DB 13; Length 559;  
Best Local Similarity 100.0%; Pred. No. 0.0016;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1575 TTATGCAAAAAAAAAAAAAAAAAAAAAA 1602  
DB 513 TTATGCAAAAAAAAAAAAAAAAAAAAAA 540

## RESULT 13

US-10-170-385-468  
Sequence 468, Application US/10170385  
Publication No. US20030203372A1  
GENERAL INFORMATION:

APPLICANT: Ward, Neil Raymond  
APPLICANT: Mundy, Christopher Robert  
APPLICANT: Kan, On  
APPLICANT: Harris, Robert Alan  
APPLICANT: White, Jonathan  
APPLICANT: Binley, Katie Mary

APPLICANT: Rayner, William Nigel

APPLICANT: Naylor, Stuart

APPLICANT: Kingsman, Susan Mary

APPLICANT: Krige, David

; TITLE OF INVENTION: ANALYSIS METHOD  
; FILE REFERENCE: 532682000100  
; CURRENT APPLICATION NUMBER: US/10/170,385  
; PRIOR FILING DATE: 2002-06-12  
; PRIOR APPLICATION NUMBER: PCT/GB02/01662  
; PRIOR FILING DATE: 2002-04-08  
; PRIOR APPLICATION NUMBER: PCT/GB01/05458  
; PRIOR FILING DATE: 2001-12-10  
; NUMBER OF SEQ ID NOS: 549  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 468  
; LENGTH: 559  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-10-170-385-468

Query Match 1.7%; Score 28; DB 13; Length 559;  
Best Local Similarity 100.0%; Pred. No. 0.0016;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1575 TTATGCAAAAAAAAAAAAAAAAAAAAAAAAAA 1602  
DB 513 TTATGCAAAAAAAAAAAAAAAAAAAAAAAAAA 540

RESULT 14  
US-10-437-963-87473  
; Sequence 87473; Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Mu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 87473  
; LENGTH: 809  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_86414C.1  
US-10-437-963-87473

Query Match 1.7%; Score 28; DB 17; Length 809;  
Best Local Similarity 100.0%; Pred. No. 0.0017;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1575 TTATGCAAAAAAAAAAAAAAAAAAAAAAAAAA 1602  
DB 780 TTATGCAAAAAAAAAAAAAAAAAAAAAAAAAA 807

RESULT 15  
US-10-437-963-47572/c  
; Sequence 47572; Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Mu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 47572  
; LENGTH: 1021  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_50327C.1  
US-10-437-963-47572

Query Match 1.7%; Score 28; DB 17; Length 1021;  
Best Local Similarity 100.0%; Pred. No. 0.0017;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1575 TTATGCAAAAAAAAAAAAAAAAAAAAAAAAAA 1602  
DB 36 TTATGCAAAAAAAAAAAAAAAAAAAAAAAAAA 9

Search completed: August 2, 2004, 16:27:05  
Job time : 774 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2004, 11:13:45 ; Search time 4443 Seconds  
(without alignments)  
10767.322 Million cell updates/sec

Title: US-10-031-331B-39  
Perfect score: 1602  
Sequence: 1 cacacgttgatttaaccat.....aaaaaaaaaaaaaaaaaaaa 1602

Scoring table: OLIGO NUC  
Gapop\_60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST.\*

1: em\_estba.\*

2: em\_esthum.\*

3: em\_estin.\*

4: em\_estmu.\*

5: em\_estov.\*

6: em\_estpl.\*

7: em\_estro.\*

8: em\_htc.\*

9: gb\_est1.\*

10: gb\_est2.\*

11: gb\_htc.\*

12: gb\_est3.\*

13: gb\_est4.\*

14: gb\_est5.\*

15: em\_estfun.\*

16: em\_estom.\*

17: em\_gss\_hum.\*

18: em\_gss\_inv.\*

19: em\_gss\_pln.\*

20: em\_gss\_vrt.\*

21: em\_gss\_fun.\*

22: em\_gss\_mam.\*

23: em\_gss\_mus.\*

24: em\_gss\_pro.\*

25: em\_gss\_rtd.\*

26: em\_gss\_pbg.\*

27: em\_gss\_vrl.\*

28: gb\_gss1.\*

29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	368	23.0	713..10..	BE231445	BE231445-S650815 S
2	35	2.2	552	10	BF051218
3	35	2.2	568	10	BF051675
4	35	2.2	583	14	CB256828

5	35	2.2	593	12	BM113004
6	35	2.2	602	10	AM623904
7	35	2.2	602	13	BQ114297
8	35	2.2	604	10	BE919890
9	35	2.2	605	10	AM621907
10	35	2.2	605	10	AM623370
11	35	2.2	629	10	AM649895
12	35	2.2	654	12	BG642645
13	35	2.2	722	12	BG128609
14	35	2.2	722	12	BI931842
15	32	2.0	595	14	CA847413
16	32	2.0	603	12	BUS55248
17	32	2.0	605	12	BUS562413
18	32	2.0	703	14	CB342962
19	31	1.9	458	12	BI841667
20	31	1.9	518	12	BI396184
21	31	1.9	560	12	BUS62806
22	31	1.9	560	13	B0718025
23	31	1.9	608	14	CD482486
24	30	1.9	122	9	A2721389
25	30	1.9	124	9	A1611367
26	30	1.9	126	9	A1252727
27	30	1.9	127	9	A1250129
28	30	1.9	128	9	A1802850
29	30	1.9	129	9	A131930
30	30	1.9	129	9	A1312227
31	30	1.9	129	9	A1312386
32	30	1.9	144	9	A1345222
33	30	1.9	146	9	A1345215
34	30	1.9	150	9	A1305426
35	30	1.9	150	9	A1205559
36	30	1.9	150	9	A1205559
37	30	1.9	156	14	CD801054
38	30	1.9	164	9	A1266670
39	30	1.9	232	13	BUS35860
40	30	1.9	237	14	CD642043
41	30	1.9	247	13	BQ800441
42	30	1.9	315	13	BUT80369
43	30	1.9	336	10	AM682563
44	30	1.9	340	10	BF014794
45	30	1.9	388	9	A1666802

## ALIGNMENTS

RESULT 1  
BE231445  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

BE231445  
S650815 Suaeda salsa ZAP CDNA library  
CDNA similar to c.elegans CDNA Y492B11.3, mRNA sequence.  
BE231445.1 GI:8996113  
EST.  
Suaeda maritima subsp. salsa  
Suaeda maritima subsp. salsa  
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales; Amaranthaceae; Suaeda.  
1 (bases 1 to 713)  
Zhang, L., Ma, C.L., Wang, P.P., Sun, Y.F., Zhao, Y.X. and Zhang, H.  
Expressed sequence tags from a halophyte Suaeda salsa CDNA library  
Unpublished (2000)  
Contact: Hui Zhang  
Key Laboratory of Plant Stress Research  
The Biology Department of Shandong Normal University  
No.88, Wenhua East Road, Jinan, Shandong Province, 250014, PRC  
Tel: (86)531-2960864  
Fax: (86)531-2966954  
Email: zhangh@snnu.edu.cn.  
Location/Qualifiers  
1. 713  
/organism="Suaeda maritima subsp. salsa"

```

/mol_type="mRNA"
/sub_species="salsa"
/db_xref="taxon:126914"
/dev_stages="seedling"
/clone_lib="Suada salsa ZAP cDNA library"
/notes="Organ: aerial part tissue; Vector: lambda zap;
Site 1: EcorI; Site 2: XhoI; total RNA extraction from
NaCl(400mM) treated Suada salsa by RNAGENT kit (Promega);
mRNA isolation by MESSAGEMAKER kit (GIBCO BRL);
directional cDNA synthesis (EcorI XhoI) by cDNA synthesis
kit (STRATAGENE); the ZAP express library by GigapackIII
Gold Cloning kit (STRATAGENE)"

```

```

ORIGIN
Query Match 23.0%; Score 368; DB 10; Length 713;
Best Local Similarity 99.6%; Pred. No. 3.2e-52;
Matches 468; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 241 AAAAATGTCAGTTTATGTGCTGATGATGACTTCTCCACTCTCAGTTTCCACACAT 300
DB 1 AAAAATGTCAGTTTATGTGCTGATGATGACTTCTCCACTCTCAGTTTCCACACAT 60
QY 301 TCATTGATGATGATTTCTCCAAATGGTGTCTCATGTATCTTCTGATGAAGAGTGGAA 360
DB 61 TCATTGATGATGATTTCTCCAAATGGTGTCTCATGTATCTTCTGATGAAGAGTGGAA 120
QY 361 AATTGCTTGAAGAAATGTTGAATGTTGAAGCAGGGGTTACATTTCTTCAGAGAA 420
DB 121 GAATTGCTTGAAGAAATGTTGAATGTTGAAGCAGGGGTTATATTTCTTCAGAGAA 180
QY 421 TCTGTTTCCATCAATCTGGGATCACAAGCGAAAGCAATCCACCACTACCGTGAA 480
DB 181 TCTGTTTCCATCAATCTGGGATCACAAGCGAAAGCAATCCACCACTACCGTGAA 240
QY 481 CTTAGGTTTCACTAAGGCTTCAAGAGTGTCTATTTGCAAGTGTATCTGCAAGTCT 540
DB 241 CTTAGGTTTCACTAAGGCTTCAAGAGTGTCTATTTGCAAGTGTATCTGCAAGTCT 300
QY 541 TATGAGTCTCCCTACTAGTGCAGAAATGTTATGAGCTTATGTGAGGATGATAGG 600
DB 301 TATGAGTCTCCCTACTAGTGCAGAAATGTTATGAGCTTATGTGAGGATGATAGG 360
QY 601 CAGAACAGATTAGTGTGTTGGCAAGAGTTGATCTTAAGGATGATAGGAGTTCCAG 660
DB 361 CAGAACAGATTAGTGTGTTGGCAAGAGTTGATCTTAAGGATGATAGGAGTTCCAG 420
QY 661 CGATTTCGATAGTACCCAGTACAAGTGTATAGCATTTCCGATATGA 710
DB 421 CGATTTCGATAGTACCCAGTACAAGTGTATAGCATTTCCGATATGA 470

```

```

RESULT 2
BF051218 552 bp mRNA linear EST 18-MAY-2001
LOCUS EST43693 tomato developing/immature green fruit Lycopersicon
DEFINITION esculentum cDNA clone cLEM21N18 5' sequence, mRNA sequence.
ACCESSION BF051218
VERSION BF051218.1 GI:10805114
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 552)
Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.B.,
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,
Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and
Tankalev,S.D.
Generation of ESTs from tomato fruit-tissue, immature-green...
Unpublished (2000)
Contact: CUGI
Clemson University Genomics Institute

```

```

Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.

FEATURES
Source
1..552
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEM21N18"
/tissue_type="fruit"
/dev_stages="immature green (5-35 days post-anthesis)"
/lab_host="SOLR"
/clone_lib="tomato developing/immature green fruit"
/notes="Vector: pBluescriptSKmCquadap; Site 1: EcorI;
Site 2: XhoI; fruit were tagged at 5 dpa (0.5 cm) and
harvested at 7 day intervals through 35 dpa. Equal masses
of tissue from each stage were combined (including seeds
and locules) prior to mRNA isolation."

```

```

ORIGIN
Query Match 2.2%; Score 35; DB 10; Length 552;
Best Local Similarity 100.0%; Pred. No. 2.3e+03;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 AATGTCACGTTTATGTGCTGATGATGACTTCTCC 278
DB 211 AATGTCACGTTTATGTGCTGATGATGACTTCTCC 245

```

```

RESULT 3
BF051675 568 bp mRNA linear EST 18-MAY-2001
LOCUS EST43692 tomato developing/immature green fruit Lycopersicon
DEFINITION esculentum cDNA clone cLEM23J8 5' sequence, mRNA sequence.
ACCESSION BF051675
VERSION BF051675.1 GI:10805571
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 568)
Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.B.,
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,
Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and
Tankalev,S.D.
Generation of ESTs from tomato fruit tissue, immature green
Unpublished (2000)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.

```

```

FEATURES
Source
1..568
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEM23J8"
/tissue_type="fruit"
/dev_stages="immature green (5-35 days post-anthesis)"
/lab_host="SOLR"
/clone_lib="tomato developing/immature green fruit"
/notes="Vector: pBluescriptSKmCquadap; Site 1: EcorI;
Site 2: XhoI; Fruit were tagged at 5 dpa (0.5 cm) and
harvested at 7 day intervals through 35 dpa. Equal masses
of tissue from each stage were combined (including seeds
and locules) prior to mRNA isolation."

```

ORIGIN

```

Query Match      2.2%; Score 35; DB 10; Length 568;
Best Local Similarity 100.0%; Pred. No. 2.2e+03;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 AATGTCAGTTATGTCGCTGATGCTGACTTCTCC 278
      |||||
DB 487 AATGTCAGTTATGTCGCTGATGCTGACTTCTCC 521

RESULT 4
LOCUS CB256828
DEFINITION 58-E012740-027-002-D16-T7R MP12-ADIS-027 Arabidopsis thaliana cDNA
ACCESSION CB256828
VERSION CB256828.1 GI:32881601
KEYWORDS EST
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 583)
AUTHORS Schmid,K.J., Soerensen,T.R., Stracke,R., Torjek,O., Altmann,T.,
Mitchell-Olds,T. and Weishaar,B.
TITLE Large-scale identification and analysis of genome-wide
single-nucleotide polymorphisms for mapping in Arabidopsis thaliana
JOURNAL Genome Res. 13 (6), 1250-1257 (2003)
MEDLINE 22683290
PUBMED 12799357
COMMENT Contact: Weishaar B
ADIS DNA core facility at MP12
Max-Planck-Institute for Plant Breeding Research
Carl-von-Linne Weg 10, 50829 Koeln, Germany
Fax: 00492215062851
Email: weishaar@mpiz-koeln.mpg.de
Insert length: 583 Std Error: 0.00
Plate: 2 row: D column: 16
Seq primer: T7R; CTAATACGACTCTACTATAGGGA.

FEATURES
source
1..583
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/cultivar="Eifel-2 (Si-2)"
/db_xref="GABI:593271"
/db_xref="taxon:3702"
/clone="MP12p772D162Q"
/tissue_type="whole plant"
/dev_stages="adult plant, mixed stresses"
/lab_host="S. coli TOP10"
/clone_lib="MP12-ADIS-027"
/notes="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; cDNA
library from Arabidopsis thaliana, accession Eifel-2; ten
week old total plants grown under long-day conditions in
soil, whole adult plants were treated for 24 hours with
different stresses, (1) at 4M-0 C in the dark, (2), at 37
Grad C in the dark, (3) lying in the lab after removing
from soil, (4) in the greenhouse after wounding leaves
with a forceps, (5) in the lab watering with a 150 mM
NaCl solution, (6) at 26 M-0C in the light/UV; equal
quantities of stressed plant material were pooled; library
was made at the Max-Planck-Institute for Plant Breeding
Research, Cologne, Germany; cloning sites SalI-NotI,
primer sites and orientation:
T7-SalI-CCAGCGTCGCG-5Prime-cDNA-polyA-CC-NotI-SP6; Note:
Sequencing granted in the context of the GABI Arabidopsis
Verbund I: Genetic Diversity, 'Establishment of
high-efficiency SNP-based mapping tools and development of
methods for genome-wide mutation detection' PI: Bernd
Weishaar Sequence submission-managed-by-RZPD/GABI-Primary
database: http://gabi.rzpd.de. This clone is available
from RZPD; contact RZPD (clone@rzpd.de) for further
information."

```

```

ORIGIN
Query Match      2.2%; Score 35; DB 14; Length 583;
Best Local Similarity 100.0%; Pred. No. 2.2e+03;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 AAGCTTCTGATCTTGACAAAGAGAACGCTCTGAG 78
      |||||
DB 313 AAGCTTCTGATCTTGACAAAGAGAACGCTCTGAG 347

RESULT 5
LOCUS BM113004
DEFINITION EST560540 potato roots Solanum tuberosum cDNA clone CPRO17C13 5'
end, mRNA sequence.
ACCESSION BM113004
VERSION BM113004.1 GI:17076052
KEYWORDS EST
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 593)
AUTHORS van der Hoeven,R., Sun,H., Karamycheva,S.A., Tsai,J., Van Aken,S.,
Utterback,T., Chiemingo,A., Bougri,O., Buell,C.R., Romning,C.,
Tankalev,S. and Baker,B.
TITLE Generation of ESTs from potato roots
JOURNAL Unpublished (2001)
COMMENT Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: T3

FEATURES
source
1..593
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="CPRO17C13"
/tissue_type="roots"
/dev_stages="in vitro grown stem cuttings"
/lab_host="SOLR"
/clone_lib="potato roots"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Cornell University, Tanksley lab;
sequencing: The Institute for Genomic Research. Roots were
isolated from in vitro grown stem cuttings on CM medium.
Roots were isolated two weeks after placing the stem
cuttings from in vitro grown plants on medium."

ORIGIN
Query Match      2.2%; Score 35; DB 12; Length 593;
Best Local Similarity 100.0%; Pred. No. 2.1e+03;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 AATGTCAGTTATGTCGCTGATGCTGACTTCTCC 278
      |||||
DB 475 AATGTCAGTTATGTCGCTGATGCTGACTTCTCC 509

RESULT 6
LOCUS AW623904
DEFINITION EST321849 tomato flower buds 3-8 mm, Cornell University
Lycopersicon esculentum cDNA clone CTOB13B5 5', mRNA sequence.
ACCESSION AW623904
VERSION AW623904.1 GI:7336931

```

## ORIGIN



Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 244 AATGTCACGTTATGTCGTCGATGACTTCCTCC 278  
 Db 500 AATGTCACGTTATGTCGTCGATGACTTCCTCC 534

RESULT 9  
 AW621907 605 bp mRNA linear EST 18-MAY-2001  
 LOCUS EST312705 tomato root during/after fruit set, Cornell University  
 DEFINITION Lycopersicon esculentum cDNA clone cLEX13N23 5', mRNA sequence.  
 AW621907  
 VERSION AW621907.1 GI:7333554  
 SOURCE EST.  
 ORGANISM Lycopersicon esculentum (tomato)  
 Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
 1 (bases 1 to 605)  
 van der Hoeven, R.S., Garvin, D., Matern, A.L., Holt, I.E., Kochian, L.,  
 Liang, F., Hansen, T., Craven, M.B., Bowman, C.L., Romning, C.M.,  
 Nierman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and  
 Tanksley, S.D.  
 Generation of ESTs from tomato root, during and after fruit set  
 Unpublished (1999)  
 CONTACT: CUGI  
 COMMENT Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 5 prime sequence.

FEATURES  
 source  
 1..605  
 Location/Qualifiers  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="cLEX13N23"  
 /tissue\_type="root"  
 /dev\_stage="plants during and after fruit set"  
 /clone\_lib="tomato root during/after fruit set, Cornell University"  
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Tanksley; Tissue supplied by Dave Garvin (USDA-ARS, Ithaca, NY 14850)."

ORIGIN  
 Query Match 2.2%; Score 35; DB 10; Length 605;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+03;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 244 AATGTCACGTTATGTCGTCGATGACTTCCTCC 278  
 Db 498 AATGTCACGTTATGTCGTCGATGACTTCCTCC 532

RESULT 10  
 AW623370 605 bp mRNA linear EST 18-MAY-2001  
 LOCUS EST312315 tomato flower buds 3-8 mm, Cornell University  
 DEFINITION Lycopersicon esculentum cDNA clone cTOB10A19 5', mRNA sequence.  
 AW623370  
 VERSION AW623370.1 GI:7336397  
 SOURCE EST.  
 ORGANISM Lycopersicon esculentum (tomato)  
 Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
 1 (bases 1 to 605)  
 van der Hoeven, R.S., Bezzereades, J.L., Matern, A.L., Holt, I.E.,

Qy 244 AATGTCACGTTATGTCGTCGATGACTTCCTCC 278  
 Db 498 AATGTCACGTTATGTCGTCGATGACTTCCTCC 532

RESULT 11  
 AW649895 629 bp mRNA linear EST 18-MAY-2001  
 LOCUS EST328349 tomato germinating seedlings, TAMU Lycopersicon  
 DEFINITION esculentum cDNA clone cLE11C11 5', mRNA sequence.  
 AW649895  
 VERSION AW649895.1 GI:7411133  
 SOURCE EST.  
 ORGANISM Lycopersicon esculentum (tomato)  
 Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
 1 (bases 1 to 629)  
 Liang, F., Hansen, T., Craven, M.B., Bowman, C.L., Romning, C.M.,  
 Nierman, W., Fraser, C.M., Giovannoni, J.J., Martin, G.B. and  
 Tanksley, S.D.  
 Generation of ESTs from germinating tomato seed  
 Unpublished (2000)  
 CONTACT: CUGI  
 COMMENT Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 5 prime sequence.

FEATURES  
 source  
 1..629  
 Location/Qualifiers  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="cLE11C11"  
 /tissue\_type="whole seedlings"  
 /dev\_stage="7 days post imbibition"  
 /clone\_lib="tomato germinating seedlings, TAMU"

ORIGIN  
 Query Match 2.2%; Score 35; DB 10; Length 605;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+03;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 244 AATGTCACGTTATGTCGTCGATGACTTCCTCC 278  
 Db 498 AATGTCACGTTATGTCGTCGATGACTTCCTCC 532

Qy 244 AATGTCACGTTATGTCGTCGATGACTTCCTCC 278  
 Db 500 AATGTCACGTTATGTCGTCGATGACTTCCTCC 534

FEATURES  
 source  
 1..605  
 Location/Qualifiers  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="cTOB10A19"  
 /tissue\_type="flower"  
 /dev\_stage="3-8mm buds"  
 /clone\_lib="tomato flower buds 3-8 mm, Cornell University"  
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Tanksley; Flower buds and flowers were taken from greenhouse plants (4-8 wks old, TA496). They were immediately frozen in liquid nitrogen and then size-separated while remaining frozen."

ORIGIN  
 Query Match 2.2%; Score 35; DB 10; Length 605;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+03;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 244 AATGTCACGTTATGTCGTCGATGACTTCCTCC 278  
 Db 498 AATGTCACGTTATGTCGTCGATGACTTCCTCC 532

RESULT 11  
 AW649895 629 bp mRNA linear EST 18-MAY-2001  
 LOCUS EST328349 tomato germinating seedlings, TAMU Lycopersicon  
 DEFINITION esculentum cDNA clone cLE11C11 5', mRNA sequence.  
 AW649895  
 VERSION AW649895.1 GI:7411133  
 SOURCE EST.  
 ORGANISM Lycopersicon esculentum (tomato)  
 Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
 1 (bases 1 to 629)  
 Liang, F., Hansen, T., Craven, M.B., Bowman, C.L., Romning, C.M.,  
 Nierman, W., Fraser, C.M., Giovannoni, J.J., Martin, G.B. and  
 Tanksley, S.D.  
 Generation of ESTs from germinating tomato seed  
 Unpublished (2000)  
 CONTACT: CUGI  
 COMMENT Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 5 prime sequence.

FEATURES  
 source  
 1..629  
 Location/Qualifiers  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="cLE11C11"  
 /tissue\_type="whole seedlings"  
 /dev\_stage="7 days post imbibition"  
 /clone\_lib="tomato germinating seedlings, TAMU"

ORIGIN  
 Query Match 2.2%; Score 35; DB 10; Length 605;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+03;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 244 AATGTCACGTTATGTCGTCGATGACTTCCTCC 278  
 Db 498 AATGTCACGTTATGTCGTCGATGACTTCCTCC 532

RESULT 11  
 AW649895 629 bp mRNA linear EST 18-MAY-2001  
 LOCUS EST328349 tomato germinating seedlings, TAMU Lycopersicon  
 DEFINITION esculentum cDNA clone cLE11C11 5', mRNA sequence.  
 AW649895  
 VERSION AW649895.1 GI:7411133  
 SOURCE EST.  
 ORGANISM Lycopersicon esculentum (tomato)  
 Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
 1 (bases 1 to 629)  
 Liang, F., Hansen, T., Craven, M.B., Bowman, C.L., Romning, C.M.,  
 Nierman, W., Fraser, C.M., Giovannoni, J.J., Martin, G.B. and  
 Tanksley, S.D.  
 Generation of ESTs from germinating tomato seed  
 Unpublished (2000)  
 CONTACT: CUGI  
 COMMENT Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 5 prime sequence.

FEATURES  
 source  
 1..629  
 Location/Qualifiers  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="cLE11C11"  
 /tissue\_type="whole seedlings"  
 /dev\_stage="7 days post imbibition"  
 /clone\_lib="tomato germinating seedlings, TAMU"

ORIGIN  
 Query Match 2.2%; Score 35; DB 10; Length 605;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+03;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 244 AATGTCACGTTATGTCGTCGATGACTTCCTCC 278  
 Db 498 AATGTCACGTTATGTCGTCGATGACTTCCTCC 532

RESULT 11  
 AW649895 629 bp mRNA linear EST 18-MAY-2001  
 LOCUS EST328349 tomato germinating seedlings, TAMU Lycopersicon  
 DEFINITION esculentum cDNA clone cLE11C11 5', mRNA sequence.  
 AW649895  
 VERSION AW649895.1 GI:7411133  
 SOURCE EST.  
 ORGANISM Lycopersicon esculentum (tomato)  
 Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
 1 (bases 1 to 629)  
 Liang, F., Hansen, T., Craven, M.B., Bowman, C.L., Romning, C.M.,  
 Nierman, W., Fraser, C.M., Giovannoni, J.J., Martin, G.B. and  
 Tanksley, S.D.  
 Generation of ESTs from germinating tomato seed  
 Unpublished (2000)  
 CONTACT: CUGI  
 COMMENT Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 5 prime sequence.

FEATURES  
 source  
 1..629  
 Location/Qualifiers  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="cLE11C11"  
 /tissue\_type="whole seedlings"  
 /dev\_stage="7 days post imbibition"  
 /clone\_lib="tomato germinating seedlings, TAMU"

ORIGIN  
 Query Match 2.2%; Score 35; DB 10; Length 605;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+03;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 244 AATGTCACGTTATGTCGTCGATGACTTCCTCC 278  
 Db 498 AATGTCACGTTATGTCGTCGATGACTTCCTCC 532

RESULT 11  
 AW649895 629 bp mRNA linear EST 18-MAY-2001  
 LOCUS EST328349 tomato germinating seedlings, TAMU Lycopersicon  
 DEFINITION esculentum cDNA clone cLE11C11 5', mRNA sequence.  
 AW649895  
 VERSION AW649895.1 GI:7411133  
 SOURCE EST.  
 ORGANISM Lycopersicon esculentum (tomato)  
 Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
 1 (bases 1 to 629)  
 Liang, F., Hansen, T., Craven, M.B., Bowman, C.L., Romning, C.M.,  
 Nierman, W., Fraser, C.M., Giovannoni, J.J., Martin, G.B. and  
 Tanksley, S.D.  
 Generation of ESTs from germinating tomato seed  
 Unpublished (2000)  
 CONTACT: CUGI  
 COMMENT Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 5 prime sequence.

FEATURES  
 source  
 1..629  
 Location/Qualifiers  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="cLE11C11"  
 /tissue\_type="whole seedlings"  
 /dev\_stage="7 days post imbibition"  
 /clone\_lib="tomato germinating seedlings, TAMU"

/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; 7 days post imbibition on water-agar. Mixed stage whole germinating seedlings from seed coat emergence up to two centimeters in seeds not showing obvious signs of germination were discarded."

## ORIGIN

Query Match 2.2%; Score 35; DB 10; Length 629;  
Best Local Similarity 100.0%; Pred. No. 2e+03;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 244 AATGTCAGTTTATGTGCTGATGACTTCTCC 278  
|||||  
DB 482 AATGTCAGTTTATGTGCTGATGACTTCTCC 516  
|||||

## RESULT 12

BG642645 654 bp mRNA linear EST 24-APR-2001  
LOCUS EST510839 tomato shoot/meristem Lycopersicon esculentum cDNA clone  
DEFINITION CTOF25C16 5' sequence, mRNA sequence.

## ACCESSION

BG642645 1 GI:13777498

## KEYWORDS

EST

## SOURCE

Lycopersicon esculentum (tomato)

## ORGANISM

Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

## REFERENCE

1 (bases 1 to 654)  
van der Hoeven, R., Bezzerides, J., Sun, H., Cho, J., Utterback, T.,

## AUTHORS

Hansen, C., Ronning, C. and Tanksley, S.

## TITLE

Generation of ESTs from tomato shoot/meristem tissue

## JOURNAL

Unpublished (2001)

## COMMENT

Contact: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>.

## FEATURES

source

1..654  
Location/Qualifiers

/organism="Lycopersicon esculentum"

/mol\_type="mRNA"

/cultivar="TA496"

/db\_xref="taxon:4081"

/clone="CTOF25C16"

/tissue\_type="shoot/meristem"

/dev\_stage="developing shoots from 4-6wks old plants"

/lab\_host="SOLR"

/clone\_lib="tomato shoot/meristem"

/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; Small expanding leaves from the growing tip were

taken from greenhouse plants (4-6wks old TA496). Tissue

was immediately frozen in liquid nitrogen."

## ORIGIN

Query Match 2.2%; Score 35; DB 12; Length 654;  
Best Local Similarity 100.0%; Pred. No. 1.9e+03;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 244 AATGTCAGTTTATGTGCTGATGACTTCTCC 278  
|||||  
DB 318 AATGTCAGTTTATGTGCTGATGACTTCTCC 352  
|||||

## RESULT 13

BG128609 722 bp mRNA linear EST 31-JAN-2001  
LOCUS EST474255 tomato shoot/meristem Lycopersicon esculentum cDNA clone  
DEFINITION CTOF21D11 5' sequence, mRNA sequence.

## ACCESSION

BG128609

## KEYWORDS

EST

## SOURCE

## ORGANISM

Lycopersicon esculentum (tomato)  
Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

## REFERENCE

## AUTHORS

1 (bases 1 to 722)  
van der Hoeven, R., Bezzerides, J., Sun, H., Cho, J., Utterback, T.,

## TITLE

Generation of ESTs from tomato shoot/meristem tissue

## JOURNAL

Unpublished (2001)

## COMMENT

Contact: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>.

## FEATURES

source

1..722  
Location/Qualifiers

/organism="Lycopersicon esculentum"

/mol\_type="mRNA"

/cultivar="TA496"

/db\_xref="taxon:4081"

/clone="CTOF21D11"

/tissue\_type="shoot/meristem"

/dev\_stage="developing shoots from 4-6wks old plants"

/lab\_host="SOLR"

/clone\_lib="tomato shoot/meristem"

/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; Small expanding leaves from the growing tip were

taken from greenhouse plants (4-6wks old TA496). Tissue

was immediately frozen in liquid nitrogen."

## ORIGIN

Query Match 2.2%; Score 35; DB 12; Length 722;  
Best Local Similarity 100.0%; Pred. No. 1.8e+03;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 244 AATGTCAGTTTATGTGCTGATGACTTCTCC 278  
|||||

DB 481 AATGTCAGTTTATGTGCTGATGACTTCTCC 515  
|||||

## RESULT 14

## BI931842

## LOCUS

## DEFINITION

EST551731 tomato flower, 8 mm to preanthesis buds Lycopersicon

esculentum cDNA clone CTOC21A20 5' end, mRNA sequence.

## ACCESSION

BI931842

## VERSION

EST.

## KEYWORDS

EST.

## SOURCE

Lycopersicon esculentum (tomato)

## ORGANISM

Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

## REFERENCE

## AUTHORS

1 (bases 1 to 761)  
van der Hoeven, R.S., Bezzerides, J.L., Karamycheva, S.A., Tsai, J.,

Utterback, T., Van Aken, S., Ronning, C.M., Nieman, W., Fraser, C.M.,

Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.

Generation of ESTs from tomato flower tissue, buds 8 mm -

preanthesis

Unpublished (2001)

Contact: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

This clone is available through the Clemson University Genomics

Institute

Seq primer: T3

Location/Qualifiers

1..761

/organism="Lycopersicon esculentum"

/mol\_type="mRNA"

Search completed: August 2, 2004, 16:05:26  
Job time : 4449 secs

/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="CTOC21A20"  
/tissue\_type="flower"  
/dev\_stage="buds 8mm to preanthesis"  
/clone\_lib="tomato flower, 8 mm to preanthesis buds"  
/notes="Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2:  
XhoI; supplier: Cornell University; sequencing: The  
Institute for Genomic Research; Flower buds and flowers  
were taken from greenhouse plants (4-8 wks old, TA496).  
They were immediately frozen in liquid nitrogen and then  
size-separated while remaining frozen."

## ORIGIN

Query Match 2.2%; Score 35; DB 12; Length 761;  
Best Local Similarity 100.0%; Pred. No. 1.7e+03;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 AATGTCAGTTTATGTCGTCTGATGACTTCTCC 278  
DB 493 AATGTCAGTTTATGTCGTCTGATGACTTCTCC 527

## RESULT 15

CA847413  
LOCUS  
DEFINITION EST0320 CM334 Root cDNA Capsicum annuum linear EST 16-DEC-2002  
Lycopersicon esculentum phosphoethanolamine N-methyltransferase  
protein, mRNA sequence.

## ACCESSION

CA847413

## VERSION

CA847413.1 GI:26997305

## KEYWORDS

EST.

## SOURCE

Capsicum annuum

## ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamiales; Solanales; Solanaceae; Capsicum.  
1 (bases 1 to 595)  
Richins R., Alvarado, K., Leger, J. and O'Connell, M.A.  
Responsive transcripts in Phytophthora capsici-challenged roots of  
Capsicum annuum

## JOURNAL

Unpublished (2002)

## COMMENT

Contact: O'Connell, MA  
Department of Agronomy and Horticulture  
New Mexico State University  
MSC 3Q, P.O. Box 30003, Las Cruces, NM 88003, USA  
Tel: 505 646 5172  
Fax: 505 646 6041  
Email: moconnel@nmsu.edu

## Seq primer: T3

## High quality sequence stop: 595.

## Location/Qualifiers

1..595  
/organism="Capsicum annuum"  
/mol\_type="mRNA"  
/strain="Criollo de Morelos 334 (CM334)"  
/db\_xref="taxon:4072"  
/clone="R8-68"  
/dev\_stage="6 hours after Phytophthora capsici  
inoculation"  
/clone\_lib="CM334 Root cDNA"  
/notes="Organ: Root; Vector: Zap; Site\_1: EcoRI; Site\_2:  
Xho 1"

## ORIGIN

Query Match 2.0%; Score 32; DB 14; Length 595;  
Best Local Similarity 100.0%; Pred. No. 6.6e+03;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 AATGTCAGTTTATGTCGTCTGATGACTTCTCC 275

DB 375 AATGTCAGTTTATGTCGTCTGATGACTTCTCC 406

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER: \_\_\_\_\_**

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**